Northwestern | Kellogg

CMS-EMS
Center for Mathematical Studies in Economics and Management Sciences

Discussion Paper #1597 – Online Supplementary Material

“Searching for “Arms”: Experimentation with Endogenous Consideration Sets”
Online Supplementary Material

Daniel Fershtman
Tel Aviv University

Alessandro Pavan
Northwestern University

May 2023

Searching for “Arms”: Experimentation with Endogenous Consideration Sets

Online Supplementary Material

Daniel Fershtman **Alessandro Pavan**

May 2023

Abstract

This document contains additional material. All sections, conditions, and results specific to this document have the suffix “S” to avoid confusion with the corresponding parts in the main text. Section S.1 establishes the optimality of the index policy claimed in part (i) of Theorem 1 in the main text by means of a novel proof that exploits the recursive characterization of the search index and the categorization of the alternatives. Section S.2 characterizes the click-through-rates (CTRs), the values-per-click (VPCs), and the purchasing probabilities, in a parametric example of the application to consumer search introduced in Section 5 in the main text. Section S.3 proves the possibility that additional ad space may be detrimental to firms' profits claimed in the main text. Section S.4 shows how the characterization of the optimal policy extends to certain problems with irreversible choice. Finally, Section S.5 discusses why an index policy need not be optimal in the presence of “meta” arms with associated super-processes.

S.1 Optimality of Index Policy

Proof of part (i) of Theorem 1 in the main text. The proof exploits the recursive representation of the search index established in part (ii) of Theorem 1, along with the representation of the DM's payoff under the index rule established in part (iii) of Theorem 1 and an appropriate description of the state space, to verify that the DM's payoff under the index policy satisfies the Bellman equation of the corresponding dynamic program.

Proof strategy. The proof is in two steps. Step 1 uses the representation of the DM's payoff under the index rule established in part (iii) of Theorem 1 in the main text 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. Step 2 then uses the results in step 1 to establish the optimality of χ^* through dynamic programming.

Step 1. In the analysis below, we find it useful to describe changes in the composition of the CS, the evolution of the search technology, as well as all information acquired about the alternatives, entirely in terms of transitions between states. Rather than keeping track of the collection of kernels $G_\xi(\vartheta^m;\mu)$ describing the conditional distributions from which the marginal signals ϑ_{m+1} are drawn, we describe directly the evolution of each alternative's state ω^P as follows. When the DM explores an alternative currently in state ω^P, its new state $\tilde{\omega}^P$ is drawn from a distribution $H_{\omega^P} \in \Delta(\Omega^P)$ that is invariant to time.[1] When the DM explores a different alternative, or expands the CS, the alternative currently in state ω^P remains in the same state with certainty at the beginning of the next period. Similarly, each time search is conducted, given the current state of the search technology ω^S, the new state of the search technology $\tilde{\omega}^S$ is drawn from a distribution $H_{\omega^S} \in \Delta(\Omega^S)$.The distributions H_{ω^s} are time-homogeneous (i.e., the evolution of the search technology depends on past search outcomes but is invariant in calendar time), and 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.

Abusing notation, then denote the state of the decision problem by a function $\mathcal{S} : \Omega \to \mathbb{N}$ that specifies, for each $\omega \in \Omega$, including $\omega \in \Omega^S$, the number of alternatives, including the search technology, that are in state ω.[2] Given this notation, for any pair of states $\mathcal{S}'$ and $\mathcal{S}''$ then define $\mathcal{S}' \vee \mathcal{S}'' \equiv (\mathcal{S}'(\omega) + \mathcal{S}''(\omega) : \omega \in \Omega)$ and $\mathcal{S}' \backslash \mathcal{S}'' \equiv (\max\{\mathcal{S}'(\omega) - \mathcal{S}''(\omega), 0\} : \omega \in \Omega)$. 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 $\mathcal{S}(\hat{\omega}^S) = 1$ and such that $\mathcal{S}(\omega^S) = 0$ 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 a state by $e(\omega)$, where $\omega \in \Omega$ is the state of the search technology in case (i) and of the single physical alternative in case (ii).[3]

Lemma S.1. *For any $v \in \mathbb{R}$ and states $\mathcal{S}'$ and $\mathcal{S}''$, $\kappa(v|\mathcal{S}' \vee \mathcal{S}'') = \kappa(v|\mathcal{S}') + \kappa(v|\mathcal{S}'')$.*

Proof of Lemma S.1. The result follows from the fact that the state of each alternative that is not explored in a given period remains unchanged, along with the fact that the time-varying components θ of the various alternatives evolve independently of one another and of the state of the search technology, given the alternatives' categories ξ. Similarly, the state of the search technology remains unchanged in periods in which search is not conducted, and evolves independently of the time-varying component θ in the state of each existing alternative, given the alternatives' categories ξ. Furthermore, the index of each alternative is a function only of the alternative's state, and the

[1] Clearly, because each alternative's category ξ is fixed, given the current state $\omega^P = (\xi, \theta)$, the distribution H_{ω^P} assigns probability one to states whose category is ξ and whose signal history $\vartheta^{m+1} = (\vartheta^m, \vartheta_{m+1})$ is a "follower" of ϑ^m, meaning that it is obtained by adding a new signal realization ϑ_{m+1} to the history ϑ^m.

[2] Clearly, with this representation, there is a unique $\hat{\omega}^s \in \Omega^S$ such that $\mathcal{S}(\omega^S) = 1$ if $\omega^s = \hat{\omega}^s$ and $\mathcal{S}(\omega^S) = 0$ 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 for all $\omega^S \in \Omega^S$, $\mathcal{S}(\omega^S) = 0$.

[3] Throughout the analysis below, we maintain the assumption that an outside option with value equal to zero is available to the DM. However, to avoid possible confusion, here we do not explicitly treat the outside option as a separate alternative.

index of search is a function only of the state of the search technology. Therefore, all indexes evolve independently of one another (conditional on the alternatives' categories), and evolve only when their corresponding decision (search or exploration of an alternative) is chosen. Since the decisions are taken under the index policy χ^*, the result follows from the fact that, starting from any state $\mathcal{S}$, the *total* time it takes to bring *all* indexes (that is, those of the alternatives in the CS as well as the index of search) below any value v is the sum (across alternatives in the CS and search) of the *individual* times necessary to bring each index below v in isolation. □

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 $\mathbb{E}\left[u|\omega^P\right]$ the immediate expected payoff from exploring an alternative in state ω^P and by $\tilde{\omega}^P$ the new state of that alternative triggered by its exploration (drawn from H_{ω^P}). Let

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

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

$$V^S(\omega^S|\mathcal{S}_0) \equiv -(1-\delta)\mathbb{E}\left[c|\omega^S\right] + \delta\mathbb{E}^{\chi^*}\left[\mathcal{V}\left(S_0 \backslash e(\omega^S) \vee e(\tilde{\omega}^S) \vee W^P(\tilde{\omega}^S)\right)|\omega^S\right] \tag{S.2}$$

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 $\mathbb{E}\left[c|\omega^S\right]$ is the immediate expected cost from searching (when the state of the search technology is ω^S), $\tilde{\omega}^S$ is the new state of the search technology, and $W^P(\tilde{\omega}^S)$ is the state of the new alternatives brought to the CS by the current search, with c and $W^P(\tilde{\omega}^S)$ jointly drawn from the distribution H_{ω^S}.[4]

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_M^A, and enlarge Ω^P to include ω_M^A. Similarly, let $e(\omega_M^A)$ 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 $\kappa(v|\mathcal{S}_0 \vee e(\omega_M^A)) = \kappa(v|\mathcal{S}_0)$, whereas if $v < M$, then $\kappa(v|\mathcal{S}_0 \vee e(\omega_M^A)) = \infty$. Hence, the representation of the DM's payoff under the index policy in part (iii) of Theorem 1 in the main text, adapted to the fictitious environment that includes the auxiliary option, implies that

$$\begin{aligned}\mathcal{V}(\mathcal{S}_0 \vee e(\omega_M^A)) &= \int_0^\infty \left(1 - \mathbb{E}^{\chi^*}\left[\delta^{\kappa(v)}|\mathcal{S}_0 \vee e(\omega_M^A)\right]\right)\mathrm{d}v = M + \int_M^\infty \left(1 - \mathbb{E}^{\chi^*}\left[\delta^{\kappa(v)}|\mathcal{S}_0\right]\right)\mathrm{d}v \\ &= \mathcal{V}(\mathcal{S}_0) + \int_0^M \mathbb{E}^{\chi^*}\left[\delta^{\kappa(v)}|\mathcal{S}_0\right]\mathrm{d}v.\end{aligned} \tag{S.3}$$

The definition of χ^*, along with Conditions (S.1) and (S.2), then imply the following:

[4]Note that $W^P(\tilde{\omega}^S)$ is a deterministic function of the new state $\tilde{\omega}^S$ of the search technology. To see this, recall that, for any $m \in \mathbb{N}$, the function E_m in the definition of the state of the search technology counts how many alternatives of each possible state ω^P have been added to the CS, as a result of the m-th search.

Lemma S.2. *For any* (ω^S, ω^P, M),

$$\mathcal{V}(e(\omega^S) \vee e(\omega^A_M)) = \begin{cases} V^S(\omega^S | e(\omega^S) \vee e(\omega^A_M)) & \text{if } M \leq \mathcal{I}^S(\omega^S) \\ M > V^S(\omega^S | e(\omega^S) \vee e(\omega^A_M)) & \text{if } M > \mathcal{I}^S(\omega^S) \end{cases} \tag{S.4}$$

$$\mathcal{V}(e(\omega^P) \vee e(\omega^A_M)) = \begin{cases} V^P(\omega^P | e(\omega^P) \vee e(\omega^A_M)) & \text{if } M \leq \mathcal{I}^P(\omega^P) \\ M > V^P(\omega^P | e(\omega^P) \vee e(\omega^A_M)) & \text{if } M > \mathcal{I}^P(\omega^P). \end{cases} \tag{S.5}$$

Proof of Lemma S.2. First note that the index corresponding to the auxiliary option is equal to M. Hence, if $M \leq \mathcal{I}^S(\omega^S)$, given $e(\omega^S) \vee e(\omega^A_M)$, χ^* prescribes to start with search, implying that $\mathcal{V}(e(\omega^S) \vee e(\omega^A_M)) = V^S(\omega^S | e(\omega^S) \vee e(\omega^A_M))$. If, instead, $M > \mathcal{I}^S(\omega^S)$, χ^* prescribes to select the auxiliary option forever, with 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 $V^S(\omega^S | e(\omega^S) \vee e(\omega^A_M)) = \mathbb{E}^{\chi^*}_{>1}\left[(1-\delta)\sum_{s=0}^{\bar{\tau}-1} \delta^s U_s + \delta^{\bar{\tau}} M | \omega^S\right]$, where $\bar{\tau}$ is the first time at which the index of search and of all the alternatives brought to the CS by search fall weakly below M, and where the expectation is under the process that obtains starting from $e(\omega^S) \vee e(\omega^A_M)$ by searching in the first period and then following the index policy in each subsequent period (the notation $\mathbb{E}^{\chi^*}_{>1}[\cdot]$ is meant to highlight that the expectation is under such a process). 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 $\mathcal{I}^S(\omega^S)$,

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

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

$$\mathbb{E}^{\chi^*}_{>1}\left[(1-\delta)\sum_{s=0}^{\bar{\tau}-1} \delta^s U_s + \delta^{\bar{\tau}} M | \omega^S\right] < M\mathbb{E}^{\chi^*}_{>1}\left[(1-\delta)\sum_{s=0}^{\bar{\tau}-1} \delta^s + \delta^{\bar{\tau}} | \omega^S\right] = M.$$

Similar arguments establish Condition (S.5). □

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 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) exploring first one of the 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.[5]

Lemma S.3. *Let* $\mathcal{S}_0$ *be the initial state of the decision problem, with* $\omega^S \in \Omega^S$ *denoting the state*

[5] In the statement of the lemma, $\mathcal{S}_0 \setminus e(\omega^S)$ is the state of a fictitious problem where search is not possible, whereas $\mathcal{S}^P_0 \setminus e(\omega^P)$ is the state of the CS obtained from $\mathcal{S}^P_0$ by subtracting an alternative in state ω^P.

of the search technology, as specified in $\mathcal{S}_0$. We have that[6]

$$D^S(\omega^S|\mathcal{S}_0) = \int_0^{\mathcal{I}^*(\mathcal{S}_0^P)} D^S(\omega^S|e(\omega^S) \vee e(\omega_v^A)) d\mathbb{E}^{\chi^*} \left[\delta^{\kappa(v)}|\mathcal{S}_0 \setminus e(\omega^S)\right] + \mathbb{E}^{\chi^*} \left[\delta^{\kappa(0)}|S_0 \backslash e(\omega^S))\right] D^S(\omega^S|e(\omega^S) \vee e(\omega_0^A)). \tag{S.6}$$

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

$$D^P(\omega^P|\mathcal{S}_0) = \int_0^{max\{\mathcal{I}^*(\mathcal{S}_0^P \backslash e(\omega^P)), \mathcal{I}^S(\omega^S)\}} D^P(\omega^P|e(\omega^P) \vee e(\omega_v^A)) d\mathbb{E}^{\chi^*} \left[\delta^{\kappa(v)}|\mathcal{S}_0 \setminus e(\omega^P)\right] + \mathbb{E}^{\chi^*} \left[\delta^{\kappa(0)}|S_0 \backslash e(\omega^P))\right] D^P(\omega^P|e(\omega^P) \vee e(\omega_0^A)). \tag{S.7}$$

Proof of Lemma S.3. Using Condition (S.3), we have that, given the state $\mathcal{S}_0 \vee e(\omega_M^A)$ of the decision problem, and $\omega^S \in \Omega^S$,

$$D^S(\omega^S|\mathcal{S}_0 \vee e(\omega_M^A)) = \mathcal{V}(\mathcal{S}_0) + \int_0^M \mathbb{E}^{\chi^*} \left[\delta^{\kappa(v)}|\mathcal{S}_0\right] \mathrm{d}v + (1-\delta)\mathbb{E}\left[c|\omega^S\right] - \delta\mathbb{E}^{\chi^*} \left[\mathcal{V}(S_0 \backslash e(\omega^S) \vee e(\tilde{\omega}^S) \vee W^P(\tilde{\omega}^S)) + \int_0^M \mathbb{E}^{\chi^*} \left[\delta^{\kappa(v)}|S_0 \backslash e(\omega^S) \vee e(\tilde{\omega}^S) \vee W^P(\tilde{\omega}^S))\right] \mathrm{d}v|\omega^S\right], \tag{S.8}$$

where the equality follows from combining (S.2) with (S.3). Similarly,

$$D^S(\omega^S|e(\omega^S) \vee e(\omega_M^A)) = \mathcal{V}(e(\omega^S)) + \int_0^M \mathbb{E}^{\chi^*} \left[\delta^{\kappa(v)}|e(\omega^S)\right] \mathrm{d}v + (1-\delta)\mathbb{E}\left[c|\omega^S\right] - \delta\mathbb{E}^{\chi^*} \left[\mathcal{V}(e(\tilde{\omega}^S) \vee W^P(\tilde{\omega}^S)) + \int_0^M \mathbb{E}^{\chi^*} \left[\delta^{\kappa(v)}|e(\tilde{\omega}^S) \vee W^P(\tilde{\omega}^S))\right] \mathrm{d}v|\omega^S\right]. \tag{S.9}$$

Differentiating (S.8) and (S.9) with respect to M, using the independence across alternatives and search and Lemma S.1, we have that

$$\frac{\partial}{\partial M} D^S(\omega^S|\mathcal{S}_0 \vee e(\omega_M^A)) = \mathbb{E}^{\chi^*} \left[\delta^{\kappa(M)}|S_0 \backslash e(\omega^S)\right] \frac{\partial}{\partial M} D^S(\omega^S|e(\omega^S) \vee e(\omega_M^A)). \tag{S.10}$$

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 indexes with value strictly higher than M, in an environment without search where the CS is the same as the one specified in $\mathcal{S}_0$. Similar arguments imply that, for any $\omega^P \in \{\hat{\omega}^P \in \Omega^P : \mathcal{S}_0(\hat{\omega}^P) > 0\}$,

$$\frac{\partial}{\partial M} D^P(\omega^P|\mathcal{S}_0 \vee e(\omega_M^A)) = \mathbb{E}^{\chi^*} \left[\delta^{\kappa(M)}|S_0 \backslash e(\omega^P)\right] \frac{\partial}{\partial M} D^P(\omega^P|e(\omega^P) \vee e(\omega_M^A)). \tag{S.11}$$

Let $M^* \equiv \max\{\mathcal{I}^*(\mathcal{S}_0^P), \mathcal{I}^S(\omega^S)\}$. Integrating (S.10) over the interval $(0, M^*)$ of possible values

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

for the auxiliary option and rearranging, we have that

$$\begin{aligned}D^S(\omega^S|\mathcal{S}_0\vee e(\omega_0^A)) &= D^S(\omega^S|\mathcal{S}_0\vee e(\omega_{M^*}^A)) - \int_0^{M^*}\mathbb{E}^{\chi^*}\left[\delta^{\kappa(v)}|S_0\backslash e(\omega^S))\right]\frac{\partial}{\partial v}D^S(\omega^S|e(\omega^S)\vee e(\omega_v^A))\mathrm{d}v\\ &= D^S(\omega^S|\mathcal{S}_0\vee e(\omega_{M^*}^A)) - D^S(\omega^S|e(\omega^S)\vee e(\omega_{M^*}^A))\\ &\quad+\mathbb{E}^{\chi^*}\left[\delta^{\kappa(0)}|S_0\backslash e(\omega^S))\right]D^S(\omega^S|e(\omega^S)\vee e(\omega_0^A))\\ &\quad+\int_0^{M^*}D^S(\omega^S|e(\omega^S)\vee e(\omega_v^A))\mathrm{d}\mathbb{E}^{\chi^*}\left[\delta^{\kappa(v)}|S_0\backslash e(\omega^S))\right],\end{aligned}$$

where the second equality follows from integration by parts and from the fact that

$$\mathbb{E}^{\chi^*}\left[\delta^{\kappa(M^*)}|S_0\backslash e(\omega^S)\right]=1.$$

That the outside option 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 $D^S(\omega^S|\mathcal{S}_0\vee e(\omega_{M^*}^A)) = D^S(\omega^S|e(\omega^S)\vee e(\omega_{M^*}^A))$.[7] Therefore, we have that

$$\begin{aligned}D^S(\omega^S|\mathcal{S}_0) &= \int_0^{M^*}D^S(\omega^S|e(\omega^S)\vee e(\omega_v^A))\mathrm{d}\mathbb{E}^{\chi^*}\left[\delta^{\kappa(v)}|S_0\backslash e(\omega^S)\right]\\ &\quad+\mathbb{E}^{\chi^*}\left[\delta^{\kappa(0)}|S_0\backslash e(\omega^S))\right]D^S(\omega^S|e(\omega^S)\vee e(\omega_0^A)).\end{aligned}\tag{S.12}$$

Similar arguments imply that

$$\begin{aligned}D^P(\omega^P|\mathcal{S}_0) &= \int_0^{M^*}D^P(\omega^P|e(\omega^P)\vee e(\omega_v^A))\mathrm{d}\mathbb{E}^{\chi^*}\left[\delta^{\kappa(v)}|S_0\backslash e(\omega^P)\right]\\ &\quad+\mathbb{E}^{\chi^*}\left[\delta^{\kappa(0)}|S_0\backslash e(\omega^P))\right]D^P(\omega^P|e(\omega^P)\vee e(\omega_0^A)).\end{aligned}\tag{S.13}$$

To complete the proof of Lemma S.3, we consider separately two cases. Case (1): given $\mathcal{S}_0$, χ^* specifies starting by exploring a physical alternative (i.e., $M^* = \mathcal{I}^*(\mathcal{S}_0^P)$). Then Condition (S.6) in the lemma follows directly from (S.12). Thus consider Condition (S.7). First observe that, for any state $\omega^P\in\Omega^P$ such that $M^*>\max\{\mathcal{I}^*(\mathcal{S}_0^P\backslash e(\omega^P)),\mathcal{I}^S(\omega^S)\}$, we have that $M^*=\mathcal{I}^P(\omega^P)$, in which case $D^P(\omega^P|\mathcal{S}_0)=D^P(\omega^P|e(\omega^P)\vee e(\omega_0^A))=0$ and the integrand $D^P(\omega^P|e(\omega^P)\vee e(\omega_v^A))$ in (S.13) is equal to zero over the interval $[0,\mathcal{I}^P(\omega^P)]$ and hence also over the interval $[0,\max\{\mathcal{I}^*(\mathcal{S}_0^P\backslash e(\omega^P)),\mathcal{I}^S(\omega^S)\}]$. We thus have that, in this case, Condition (S.7) clearly holds. Next observe that, for any state $\omega^P\in\Omega^P$ such that $M^*=\max\{\mathcal{I}^*(\mathcal{S}_0^P\backslash e(\omega^P)),\mathcal{I}^S(\omega^S)\}$, Condition (S.7) follows directly from (S.13).

Case (2): given $\mathcal{S}_0$, χ^* specifies starting with 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 (S.7) in the lemma follows directly from (S.13). That Condition (S.6) also holds follows from the fact that, in this case, $D^S(\omega^S|\mathcal{S}_0)=D^S(\omega^S|e(\omega^S)\vee e(\omega_0^A))=0$ and the integrand $D^S(\omega^S|e(\omega^S)\vee e(\omega_v^A))$ in (S.12) is equal

[7] 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 $\mathbb{E}^{\chi^*}\left[\mathcal{V}\left(\mathcal{S}_0\backslash e(\omega^S)\vee e(\tilde{\omega}^S)\vee W^P(\tilde{\omega}^S)\vee e(\omega_{M^*}^A)\right)|\omega^S\right]=\mathbb{E}^{\chi^*}\left[\mathcal{V}\left(e(\tilde{\omega}^S)\vee W^P(\tilde{\omega}^S)\vee e(\omega_{M^*}^A)\right)|\omega^S\right]$. Intuitively, under the index policy, any alternative with index strictly below M^* is never explored given the presence of an auxiliary alternative with payoff M^*.

to zero over the entire interval $\left[0, \max\{\mathcal{I}^*(\mathcal{S}_0^P \backslash e(\omega^P)), \mathcal{I}^S(\omega^S)\}\right]$. □

Step 2. Using the characterization of the payoff differentials in Lemma S.3, we now establish that the average per-period payoff under χ^* solves the Bellman equation for our dynamic optimization problem. 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 S.4. *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. *$\mathcal{V}(\mathcal{S}_0) \geq V^S(\omega^S|\mathcal{S}_0)$, and $\mathcal{V}(\mathcal{S}_0) = V^S(\omega^S|\mathcal{S}_0)$ if and only if $\mathcal{I}^S(\omega^S) \geq \mathcal{I}^*(\mathcal{S}_0^P)$;*

2. *for any $\omega^P \in \{\hat{\omega}^P \in \Omega^P : \mathcal{S}_0(\hat{\omega}^P) > 0\}$, $\mathcal{V}(\mathcal{S}_0) \geq V^P(\omega^P|\mathcal{S}_0)$, and $\mathcal{V}(\mathcal{S}_0) = V^P(\omega^P|\mathcal{S}_0)$ 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 S.4. *Part 1.* First, use (S.4) to note that, for all $v \geq 0$, $D^S(\omega^S|e(\omega^S) \vee e(\omega_v^A)) \geq 0$, with the inequality holding as an equality if and only $v \leq \mathcal{I}^S(\omega^S)$. Therefore, from (S.6), $D^S(\omega^S|\mathcal{S}_0) \geq 0$ – and hence $\mathcal{V}(\mathcal{S}_0) \geq V^S(\omega^S|\mathcal{S}_0)$ – with the inequality holding as an equality if and only if $\mathcal{I}^*(\mathcal{S}_0^P) \leq \mathcal{I}^S(\omega^S)$.

Part 2. Similarly, use (S.5) to observe that for any $\omega^P \in \{\hat{\omega}^P \in \Omega^P : \mathcal{S}_0^P(\hat{\omega}^P) > 0\}$ and any $v \geq 0$, $D^P(\omega^P|e(\omega^P) \vee e(\omega_v^A)) \geq 0$, with the inequality holding as an equality if and only if $0 \leq v \leq \mathcal{I}^P(\omega^P)$. Therefore, from (S.7), $D^P(\omega^P|\mathcal{S}_0) \geq 0$ with the inequality holding as equality if and only if $\mathcal{I}^P(\omega^P) \geq \max\{\mathcal{I}^*(\mathcal{S}_0^P \backslash e(\omega^P)), \mathcal{I}^S(\omega^S)\}$. The result in part 2 then follows from the fact that the last inequality holds if and only if $\mathcal{I}^P(\omega^P) = \mathcal{I}^*(\mathcal{S}_0^P) \geq \mathcal{I}^S(\omega^S)$.

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 that $\mathcal{V}(\mathcal{S}_0) = \mathcal{V}^*(\mathcal{S}_0)$, and hence the optimality of the index policyχ^*. □

This completes the proof. ■

S.2 CTRs, VPCs and eventual purchase probabilities: a parametric example

Consider a market with two firms, each of which advertises a single product. Let z denote the profit a firm derives from selling its product and assume that the two firms' profits are drawn independently from a distribution Z. Suppose that the product of each firm can either be highly attractive to the consumer ($\xi = H$) or less attractive ($\xi = L$), with the types drawn independently from $\Xi = \{H, L\}$, with $Pr(\xi = H) = q^H$. A highly-attractive product yields an utility to the consumer, net of the purchasing price, drawn uniformly from $[0, 1+\alpha]$, where $\alpha > 0$. A less-attractive product, instead, yields a net utility drawn uniformly from $[0,1]$. The consumer learns the attractiveness of a firm's product by reading the firm's ad but discovers her net value for the product only by clicking on the ad and being directed to the firm's webpage. For simplicity, assume that $\lambda \equiv c \equiv 0$, so that the only cost is discounting.

The two firms advertise their products on a platform using the ascending-clock version of the generalized second-price auction to allocate the two ad positions. The firm dropping out first is allocated the second position and pays nothing, whereas the other firm is allocated the first position and pays to the platform the price at which the other firm dropped out per click.

Using the formula for the reservation prize in Proposition 2 in the main text characterizing the optimal policy in Pandora's-boxes problem with an endogenous CS, the "clicking index" of an L-type firm is given by

$$\mathcal{I}_L^P \equiv \mathcal{I}^P(L,\emptyset) = \frac{\frac{\delta}{2(1-\delta)^2}\left((1-\delta)^2 - (\mathcal{I}^P(L,\emptyset))^2\right)}{1+\frac{\delta}{(1-\delta)^2}\left((1-\delta) - \mathcal{I}^P(L,\emptyset)\right)},$$

which, solving for $\mathcal{I}^P(L,\emptyset)$, yields

$$\mathcal{I}_L^P = \frac{1-\delta}{\delta}\left(1-\sqrt{1-\delta^2}\right).$$

Similarly, the clicking index of an H-type firm is

$$\mathcal{I}_H^P \equiv \mathcal{I}^P(H,\emptyset) = (1+\alpha)\left(\frac{1-\delta}{\delta}\right)\left(1-\sqrt{1-\delta^2}\right).$$

Because the consumer always reads the first ad (as there are no direct cost of reading and the outside option is equal to 0), the initial search index (corresponding to the decision to read the first ad) plays no role in the analysis and hence we do not provide its characterization. The following lemma characterizes in closed form the index for reading the second ad which, which is a function of ρ^L and ρ^H. The probabilities ρ^L and ρ^H, may of course depend on the type of ad encountered in the first position.

Lemma S.5. *Let ρ^L and ρ^H represent the probabilities that the consumer assigns to finding an L or H firm in the second position. The index $\mathcal{I}^S$ for the decision to read the second ad is equal to*

$$\mathcal{I}^S(\rho^L,\rho^H) = \frac{(1-\delta)(1+\alpha)\left[1-\sqrt{1-\frac{\delta^4}{1+\alpha}\left(1+\alpha\rho^L\right)\left(1+\alpha\rho^H\right)}\right]}{\delta^2(1+\alpha\rho^L)}$$

if

$$\frac{\delta\left(1-\sqrt{1-\delta^2}\right)\left(1+\alpha\rho^L\right)}{1+\alpha} > 1-\sqrt{1-\frac{\delta^4}{1+\alpha}\left(1+\alpha\rho^L\right)\left(1+\alpha\rho^H\right)}$$

and otherwise,

$$\mathcal{I}^S(\rho^L,\rho^H) = \frac{(1-\delta)(1+\alpha)\left[1-\delta+\delta\rho^H-\sqrt{(1-\delta+\delta\rho^H)^2-\delta^4\left(\rho^H\right)^2}\right]}{\delta^2\rho^H}.$$

Proof of Lemma S.5. Recall that the index of search in the Pandora's boxes problem with endogenous CS is given by

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

where $\Xi(l)\equiv\left\{\xi\in\Xi:\ \mathcal{I}^P(\xi,\emptyset)>l\right\}$. Since there are no direct costs,

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

Note that since $\mathcal{I}^P(\emptyset,H)>\mathcal{I}^P(\emptyset,L)$, it cannot be that $\Xi(\mathcal{I}^S(\rho^L,\rho^H))=\{L\}$. Furthermore, it cannot be that $\Xi(\mathcal{I}^S(\rho^L,\rho^H))=\emptyset$, as in this case no category has an index greater than search, which means stopping (in the definition of search index) occurs immediately after search is carried out, yielding $\mathcal{I}^S(2)=0$, a contradicting. Hence, there are two feasible cases to consider: (i) $\Xi(\mathcal{I}^S(\rho^L,\rho^H))=\{H,L\}$ and (ii) $\Xi(\mathcal{I}^S(\rho^L,\rho^H))=\{H\}$. We derive the index of search for each of these cases.

Denote $\overline{\xi}\equiv\rho^L+\rho^H(1+\alpha)$, $\hat{\xi}\equiv\rho^L+\frac{\rho^H}{1+\alpha}$.

Case (i) - $\Xi(\mathcal{I}^S(\rho^L,\rho^H))=\{H,L\}$. In this case, the index of search is

$$\mathcal{I}^S(\rho^L,\rho^H) = \frac{-\delta\left[\rho^L\left(\delta\int_{\frac{\mathcal{I}^S(m)}{1-\delta}}^{\infty}v\mathrm{d}F^L(u)\right)+\rho^H\left(\delta\int_{\frac{\mathcal{I}^S(m)}{1-\delta}}^{\infty}v\mathrm{d}F^H(u)\right)\right]}{1+\left(\rho^L\left[\delta+\frac{\delta^2}{1-\delta}\left(1-F^L\left(\frac{\mathcal{I}^S(m)}{1-\delta}\right)\right)\right]+\rho^H\left[\delta+\frac{\delta^2}{1-\delta}\left(1-F^H\left(\frac{\mathcal{I}^S(m)}{1-\delta}\right)\right)\right]\right)},$$

which in the current example, after some algebra, can be rewritten as

$$\mathcal{I}^S(\rho^L,\rho^H) = \frac{-(1-\delta)^2\left(-\frac{\delta^2}{2}\overline{\xi}\right)-\frac{\delta^2}{2}\left(\mathcal{I}^S(m)\right)^2\hat{\xi}}{1-\delta-\delta^2\hat{\xi}\mathcal{I}^S(m)}.$$

Solving for $\mathcal{I}^S(\rho^L, \rho^H)$, we have that in case (i),

$$\mathcal{I}^S(\rho^L, \rho^H) = \frac{1-\delta}{\delta^2 \hat{\xi}} \left(1 - \sqrt{1 - \delta^4 \hat{\xi}\overline{\xi}}\right)$$
$$= \frac{(1-\delta)(1+\alpha)}{\delta^2 (1+\alpha\rho^L)} \left(1 - \sqrt{1 - \delta^4 \frac{(1+\alpha\rho^L)(1+\rho^H\alpha)}{1+\alpha}}\right)$$

Case (ii) - $\Xi(\mathcal{I}^S(\rho^L, \rho^H)) = \{H\}$. In this case,

$$\mathcal{I}^S(\rho^L, \rho^H) = \frac{\delta^2 \rho^H \left(\int_{\frac{\mathcal{I}^S(m)}{1-\delta}}^{\infty} v \mathrm{d}F^H(u)\right)}{1 + \rho^H \left[\delta + \frac{\delta^2}{1-\delta}\left(1 - F^H\left(\frac{\mathcal{I}^S(m)}{1-\delta}\right)\right)\right]},$$

which, after some algebra, can be written as

$$\mathcal{I}^S(m) = \frac{\frac{1}{2}\delta^2(1-\delta)^2\rho^H(1+\alpha) - \frac{\delta^2\rho^H}{2(1+\alpha)}\left(\mathcal{I}^S(m)\right)^2}{(1-\delta)^2 + \rho^H\left(\delta(1-\delta) - \delta^2\left(\frac{\mathcal{I}^S(m)}{1+\alpha}\right)\right)}.$$

Solving for $\mathcal{I}^S(\rho^L, \rho^H)$, we have that in case (ii),

$$\mathcal{I}^S(m) = \frac{(1-\delta)(1+\alpha)}{\delta^2 \rho^H}\left(1 - \delta + \delta\rho^H - \sqrt{(1-\delta+\delta\rho^H)^2 - \delta^2(\rho^H)^2}\right).$$

Now, case (i), $\Xi(\mathcal{I}^S(\rho^L, \rho^H)) = \{H, L\}$, is the relevant case if and only if

$$\mathcal{I}^P(\emptyset, L) = \frac{1-\delta}{\delta}\left(1 - \sqrt{1-\delta^2}\right) > \frac{1-\delta}{\delta^2\hat{\xi}}\left(1 - \sqrt{1-\delta^4\hat{\xi}\overline{\xi}}\right),$$

where recall that the RHS of the latter inequality is the index of search in case (i). The latter inequality can equivalently be written as

$$\delta\left(\frac{1+\alpha\rho^L}{1+\alpha}\right)\left(1 - \sqrt{1-\delta^2}\right) > 1 - \sqrt{1 - \frac{\delta^4}{1+\alpha}(1+\alpha\rho^L)(1+\alpha\rho^H)}. \tag{S.14}$$

We have therefore shown that the search index is equal to

$$\frac{(1-\delta)(1+\alpha)}{\delta^2(1+\alpha\rho^L)}\left(1 - \sqrt{1 - \frac{\delta^4}{1+\alpha}(1+\alpha\rho^L)(1+\alpha\rho^H)}\right)$$

if (S.14) holds, and otherwise it is equal to

$$\frac{(1-\delta)(1+\alpha)}{\delta^2\rho^H}\left(1 - \delta + \delta\rho^H - \sqrt{(1-\delta+\delta\rho^H)^2 - \delta^4(\rho^H)^2}\right).$$

■

Given the lemma above, the CTRs, VPCs, and eventual purchase decisions, are pinned down by the results in Section 5 in the main text.

S.3 Additional ad space - comparative statics

Consider the setting described at the end of Section 5 in the main body. The following result illustrates how the increase in the probability that search brings an additional product by firm ξ may reduce the index of search, inducing the consumer to visit the website of one of firm ξ's competitors before searching for new products. When strong enough, such an effect may reduce the probability that one of firm ξ's product is selected, and hence firm ξ's profits.

Corollary S.1. *Consider the environment described at the end of Section 5 in the main body. An increase in the probability ρ^ξ that search brings an additional product from firm ξ may reduce firm ξ's ex-ante expected profits.*

Proof of Corollary S.1. Suppose that each F^ξ is a Bernoulli distribution assigning probability p^ξ to $v = \hat{v}^\xi$ and $(1 - p^\xi)$ to $v = 0$, with $\hat{v}^\xi \in \mathbb{R}_{++}$.[8] Each firm makes equal profits on each of its two products. Hence, each firm's ex-ante total profits are equal to the total probability with which one of its two products is selected. To keep things simple, suppose the consumer incurs no cost for inspecting any product other than the time cost of postponing the final purchase: that is, $\lambda^\xi = 0$ for $\xi = A, B, C$. The consumer's discount factor is δ.

Exogenous CS. Suppose the identity of the firm receiving the additional slot is determined ex-ante, i.e., before the consumer starts the exploration process. Given the composition of the CS, the consumer then sequentially decides which product to inspect and when to stop, at which point she either chooses one of the inspected products or her outside option (whose value is normalized to zero). As shown in the main text, the reservation price for each ξ's product, before the latter is inspected, is equal to

$$\mathcal{I}(\xi, \emptyset) = \frac{(1-\delta)\delta p^\xi \hat{v}^\xi}{1 - \delta + \delta p^\xi},$$

whereas the reservation price of each ξ's product after it is inspected is equal to $\mathcal{I}(\xi, v) = (1 - \delta)v$, with $v \in \{\hat{v}^\xi, 0\}$. 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. Clearly, in this environment, each firm benefits from an increase in the probability it is given a second slot.

Endogenous CS. Now suppose that the consumer's initial CS consists of three products, one from each firm $\xi = A, B, C$, and that the CS can be expanded only once, with the expansion bringing an additional product drawn from Ξ according to ρ, with $\rho^\xi \geq 0$, $\xi = A, B, C$, and with $\sum_\xi \rho^\xi = 1$. The result in the corollary then follows from Claim S.1 below.

[8] One can think of $\hat{v}^\xi$ as the value (net of price) to the consumer in case the product is a good match, and p^ξ as the probability of such an event.

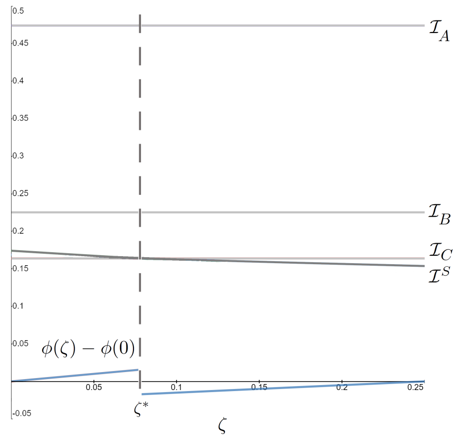

Figure 1: The change $\phi(\zeta) - \phi(0)$ in the probability with which a product of firm B is selected, as a function of ζ (in blue).

Claim S.1. Suppose that $\mathcal{I}(A, \emptyset) > \mathcal{I}(B, \emptyset) > \mathcal{I}(C, \emptyset)$. There exist parameter values consistent with the above inequalities such that an increase in ρ^B, together with a reduction by the same amount in ρ^A, leads to a decrease in the overall probability that one of firm B's products is sold (and hence in its ex-ante expected profits).

Proof. We establish the claim above by showing that an increase in the probability that search brings an extra B-product (along with a reduction by the same amount in the probability that it brings an A-product) may reduce the attractiveness of search thus inducing the consumer to inspect firm C's product before expanding the CS. We show that this effect may imply a drop in firm B's ex-ante profits.

It is easy to verify that the index for search is equal to

$$\mathcal{I}^S = \delta^2 \max_{k \in \{A,B,C\}} \left\{ \frac{\sum_{\xi \in \{\xi' \in \Xi : \mathcal{I}(\xi', \emptyset) \geq \mathcal{I}(k, \emptyset)\}} \rho^\xi p^\xi \hat{v}^\xi}{1 + \sum_{\xi \in \{\xi' \in \Xi : \mathcal{I}(\xi', \emptyset) \geq \mathcal{I}(k, \emptyset)\}} \rho^\xi \delta \left(1 + \frac{p^\xi \delta}{1-\delta}\right)} \right\}. \tag{S.15}$$

For concreteness, let $\delta = 0.9$ and suppose that $(\hat{v}^A, p^A) = (10, \frac{1}{10})$, $(\hat{v}^B, p^B) = (3, \frac{1}{3})$, and $(\hat{v}^C, p^C) = (2, \frac{1}{2})$. Note that the distributions F^ξ from which the consumer's values for the firms' products are drawn have the same mean, but are mean preserving spreads of one another; hence $\mathcal{I}(A, \emptyset) > \mathcal{I}(B, \emptyset) > \mathcal{I}(C, \emptyset)$. Suppose that, initially, $\rho^A = \rho^B = \frac{1}{4}$, and $\rho^C = \frac{1}{2}$. It is easily verified that $\mathcal{I}(A, \emptyset) = 0.473$, $\mathcal{I}(B, \emptyset) = 0.225$, $\mathcal{I}(C, \emptyset) = 0.163$, and $\mathcal{I}^S = 0.174$. Also note that $\mathcal{I}(C, \emptyset) < \mathcal{I}^S < \mathcal{I}(B, \emptyset)$, so that $\mathcal{I}^S$ does not take into account the benefits from inspecting firm C's additional product, in case search brings a second product by firm C.

Now suppose ρ^B is increased by $\zeta \in [0, 0.25]$ while ρ^A is reduced by the same amount. Let $\phi(\zeta)$ denote the probability that one of firm B's products is ultimately chosen when the probability that search brings a B-product is $\rho^B + \zeta$. Figure 1 depicts the change $\phi(\zeta) - \phi(0)$ in the probability that one of firm B's products is selected as a function of ζ, where $\phi(0) = (1 - p^A)(p^B + (1 -$

$p^B)\rho^B p^B) = 0.35$. The horizontal gray lines correspond to the indices $\mathcal{I}(A,\emptyset)$, $\mathcal{I}(B,\emptyset)$, and $\mathcal{I}(C,\emptyset)$, whereas the dark gray curve depicts $\mathcal{I}^S$, as a function of ζ. Note that $\mathcal{I}^S$ is decreasing in ζ, since $\mathcal{I}(A,\emptyset) > \mathcal{I}(B,\emptyset)$. Hence, an increase in ζ implies a lower index for search. $\mathcal{I}^S$ starts out above $\mathcal{I}(C,\emptyset)$, and intersects $\mathcal{I}(C,\emptyset)$ at an interior ζ (smaller than 0.25), denoted by ζ^* (the vertical dashed line). For $\zeta < \zeta^*$, $\mathcal{I}(C,\emptyset) < \mathcal{I}^S < \mathcal{I}(B,\emptyset)$, whereas for $\zeta > \zeta^*$, $\mathcal{I}^S < \mathcal{I}(C,\emptyset)$. The function $\mathcal{I}^S(\zeta)$ has a kink at $\zeta = \zeta^*$. For $\zeta \in [0,\zeta^*)$, the CS is expanded before firm C's product is inspected, whereas for $\zeta \in (\zeta^*, 0.25]$ the opposite is true. Therefore, the probability that one of firm B's products is chosen is equal to $\phi(\zeta) = (1-p^A)(p^B + (1-p^B)(\rho^B+\zeta)p^B)$ for $\zeta \in [0,\zeta^*)$ and is equal to $\phi(\zeta) = (1-p^A)\left(p^B + (1-p^B)(1-p^C)(\rho^B+\zeta)p^B\right)$ for $\zeta \in (\zeta^*, 0.25]$, with a downward discontinuity at $\zeta = \zeta^*$ equal to $(1-p^A)(1-p^B)p^B p^C(\rho^2+\zeta^*)$. Furthermore, the downward drop in $\phi(\zeta)$ at $\zeta = \zeta^*$ makes $\phi(\zeta) - \phi(0)$ negative over $(\zeta^*, 0.25]$, thus establishing the claim above. □

S.4 Irreversible Choice Among Alternatives

Consider the following amendment to the general model of Section 2 in the main text. At any period t, in addition to exploring an alternative in the CS or expanding the latter, the DM can *irreversibly commit* to any alternative in the CS, provided that the alternative has been explored at least M_ξ times (with ξ denoting the alternative's category).[9] Once the DM irreversibly commits to an alternative, there are no further decisions to be made. Irreversibly committing to an alternative 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. In particular, denote by $R(\omega^P)$ the *expected flow value* from irreversibly committing to an alternative whose current state is $\omega^P = (\xi,\theta)$. Note that $R(\omega^P)$ admits two equivalent interpretations: (i) the DM obtains an immediate expected payoff equal to $R(\omega^P)/(1-\delta)$ after which there are no further payoffs; (ii) payoffs continue to accrue at all subsequent periods after the irreversible choice is made, with each expected flow payoff equal to $R(\omega^P)$.

For any $\omega^P = (\xi,\theta)$ and $\hat{\omega}^P = (\hat{\xi},\hat{\theta})$, say that $\hat{\omega}^P$ "follows" ω^P if and only if $\hat{\xi} = \xi$, $\theta = (\vartheta_1, ..., \vartheta_m)$, for some m, and $\hat{\theta} = (\vartheta_1, ..., \vartheta_m, ..., \vartheta_{\hat{m}})$ for some $\hat{m} \geq m$. Denote this relation by $\hat{\omega}^P \succeq \omega^P$.

Condition 1. A category-ξ alternative has the *better-later-than-sooner property* if, for any $\omega^P = (\xi,\theta)$ such that $\theta = (\vartheta_1, ..., \vartheta_m)$, 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 S.1 (Weitzman's generalized problem). Consider the following extension of Weitzman's original problem: (i) The set of boxes is endogenous; (ii) each category-ξ box 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;

[9] If $M_\xi = 0$, the DM can irreversibly commit to any ξ-alternative without first exploring it.

(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 based on 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-th exploration and is equal to the box's prize.

Example S.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 category ξ in state $\omega^P = (\xi, \theta)$ is a function of the information $\theta = (\vartheta_1, ..., \vartheta_m)$ 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, $R(\omega^P) \leq 0$. Else, for any $\theta = (\vartheta_1, ..., \vartheta_m)$ and $\hat{\theta} = (\vartheta_1, ..., \vartheta_m, ..., \vartheta_{\hat{m}})$ such that $\hat{m} \geq m$, $R(\xi, \hat{\theta}) \geq R(\xi, \theta)$. Contrary to Weitzman's generalized problem above, the DM may derive a higher (expected) value from using an alternative without irreversibly committing to it (i.e, from leasing instead of selling) 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 alternative in state $\omega^P \in \Omega^P$ as follows:

$$\mathcal{I}^P(\omega^P) \equiv \sup_{\pi,\tau} \frac{\mathbb{E}^\pi\left[\sum_{s=0}^{\tau-1} \delta^s U_s | \omega^P\right]}{\mathbb{E}^\pi\left[\sum_{s=0}^{\tau-1} \delta^s | \omega^P\right]}, \tag{S.16}$$

where τ is a stopping time, and where π is a rule specifying whether the DM explores the alternative, or irreversibly commits to it. Similarly, modify the index of search $\mathcal{I}^S(\omega^S)$ by letting the rule π now specify not only whether the DM keeps searching or explores one of the alternatives brought to the CS through search, but also whether or not 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 alternative in the CS and the expected "retirement" value R of each alternative in the CS; (b) experiment with an 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 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 is greater than its index $\mathcal{I}^P$, as well as the index of search and both the index and the expected retirement value of any other alternative in the CS.

We then have the following result:

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

The result is established by considering a fictitious problem without irreversible choice in which, each time the DM experiments with an alternative and changes its state to 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 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 exploration of the 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.

Proof of Theorem S.1. 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 learned over time.

Consider the following *fictitious environment*, where all choices are *reversible*. Whenever an alternative of category ξ is brought to the CS, an additional *auxiliary* alternative is also introduced into the CS, yielding a fixed flow payoff of $R(\xi, \emptyset)$.[10] Furthermore, whenever a non-auxiliary alternative in state ω^P is explored, a new auxiliary alternative yielding a fixed payoff of $R(\tilde{\omega}^P)$ is also added to the CS, where $\tilde{\omega}^P$ denotes the new state of the explored alternative drawn from H_{ω^P}, as in the baseline model.[11] We say that an auxiliary alternative *corresponds to a (non-auxiliary) alternative in state ω^P* if it has been introduced to the CS as the result of either search (in which case $\theta = \emptyset$) or the exploration of an alternative in state ω^P. In this auxiliary environment, define the index of search as in the main text, with the rule π specifying whether to keep searching or exploring one of the alternatives introduced through search, including the auxiliary alternatives brought to the CS by search or by the explorations of the alternatives brought to the CS through search. For each alternative in state ω^P, define its new index as in (S.16), with the rule π in the definition of the index specifying for each period prior to stopping whether to explore the alternative itself or one of the auxiliary alternatives introduced as the result of the alternative's current and future explorations (i.e., following the period at which the index is computed; importantly, π excludes

[10] Recall that $R(\xi, \emptyset)$ is the retirement value of a physical alternative of category ξ that has never been explored.

[11] If $M_\xi > 0$, the introduction of the auxiliary alternative as the result of the exploration of an alternative in state $\omega^P = (\xi, \theta)$ occurs only if $\theta = (\vartheta_1, ..., \vartheta_s)$ with $s \geq M_\xi$, that is, only if the alternative has been explored at least M_ξ times.

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 in the main text imply that, in this auxiliary environment, the index policy based on the above new indices is optimal.[12] It 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) alternatives can be retired also in states that are not feasible any more due to the subsequent explorations of the same alternative. 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 fictitious environment. To see this, first observe that, in the fictitious environment, once the DM explores 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 fictitious environment plays no role. Next, observe that Condition 1 implies that, in the fictitious environment, if the DM selects an auxiliary alternative, she always picks the one corresponding to the "newest" state of the corresponding non-auxiliary alternative that created it; this is because the latest has the highest expected value R among all the auxiliary alternatives corresponding to the same non-auxiliary alternative. This implies that the non-perishability of the older versions of the auxiliary alternatives in the fictitious environment also plays no role. The same condition also guarantees that the policy π in the definition of the index of the non-auxiliary alternatives in the fictitious problem coincides with the one in (S.16) where the selection π is restricted to be over the exploration of the non-auxiliary 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 fictitious environment, an auxiliary alternative is introduced into the CS only when its corresponding non-auxiliary alternative has been explored more than M_ξ times, with M_ξ possibly stochastic and learned over time (in this latter case, the time-varying component of an alternative's state, θ, may also contain information about M_ξ). ■

S.5 Sub-optimality of Index Policies with "Meta-Arms"

In this section, we briefly illustrate, by means of an example, why multi-armed bandit problems in which alternatives take the form of "meta arms", i.e., sub-decision problems with their own sub-decisions, typically do not admit an index solution. This is so 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 the case even if a subset of dependent alternatives evolves independently of all other alternatives, and even if one knows how to optimally choose among the dependent alternatives in

[12] The proof must be adjusted to accommodate for the auxiliary alternatives introduced as the result of the DM exploring the physical alternatives. Since all the steps are virtually the same, the proof is omitted.

each given subset in isolation.

Consider the following extension of the environment described in the main text. There are $k \in \mathbb{N}$ sets of arms, $K_1, ..., K_k$. Arms from different sets evolve independently of one another, but the state of each arm within a set may depend on the state of other arms from the same set. More generally, suppose that each arm corresponds to a "meta arm", the activation of which involves decisions other than when to stop using it. Each meta arm has its own decision process which is independent of the other meta arms.

Clearly, the model in the main body of the paper is a special case of this richer setting. Suppose that, for each set of arms K_i, one can compute the optimal sequence of pulls, independently of the other sets of arms. Equivalently, suppose that for each "meta arm" one can compute the optimal sequence of decisions that define the usage of that arm, independently of the solution to the other meta arms' problems. It is tempting to conjecture that one may then assign an independent index to each set of arms K_i (alternatively, to each "meta arm") and that the optimal policy for the overall problem reduces to an index policy, whereby the meta arm with the highest index is selected in each period.

Perhaps surprisingly, the optimal policy for this enriched problem does not admit an index representation. When arms are not defined as in the canonical multi-armed bandit problem, but rather feature a more complicated internal decision problem (preserving the independence across arms), the optimal policy need not be an index policy. The following example illustrates.

Example S.3. There are two arms. Arm 1 yields a reward of 1,000 when it is first pulled. In all subsequent pulls, it yields a reward λ, where λ is initially unknown and may be either 1 or 10, with equal probability. After the first pull of arm 1, λ is perfectly revealed and is fully persistent. Arm 2 is a "meta arm" corresponding to the following decision problem. When the decision maker pulls arm 2 for the first time, she must also choose *how* to pull it. There are two ways to pull this arm, 2(A) and 2(B). If the decision maker selects $2(A)$, the arm yields a reward of 100 for a single period, followed by no rewards thereafter. If, instead, the decision maker selects $2(B)$, the arm yields a reward equal to 11 in each of its subsequent pulls. The choice of which version of arm 2 to use must be made the first time that arm 2 is pulled and can not be reversed.

Assume $\delta = 0.9$. The optimal policy for this problem is the following. In period 1, arm 1 is pulled. If $\lambda = 10$, then arm 2 in version $2(A)$ is then pulled for a single period, followed by arm 1 again in all subsequent periods. If, instead, $\lambda = 1$, arm 2 is then pulled in version $2(B)$ in all subsequent periods. Note that, under the optimal policy, the decision of how to use arm 2 depends on the realization of arm 1's first pull. It is then evident that the optimal policy is not an index policy, no matter how one defines the indices. This is because an index policy requires that both the index of each arm and its utilization (when an arm can be used in different versions, as in the case of "meta arm" 2 in this example) be invariant in the results of the activation of all other arms.