Searching for Arms: Experimentation with Endogenous Consideration Sets

Daniel Fershtman   Alessandro Pavan

Tel Aviv University / Northwestern University

March 2021, Princeton

What we do

- Study sequential experimentation with endogenous set of alternatives
- Alternatives come from deliberate decision to search for more options
- Tradeoff:
  - Exploring alternatives already in "consideration set" (CS)
  - Expanding CS by searching for more options

Examples

- Consumer sequentially explores products + searches for more options
- Firm evaluates candidates + expands candidate pool by searching for more
- R&D: pursuing alternative technologies + searching for new ones to explore
- Researcher alternates between ongoing projects + searches for new ideas

What we do

- Study sequential experimentation with endogenous set of alternatives
- Alternatives come from deliberate decision to search for more options
- Tradeoff:
  - Exploring alternatives already in "consideration set" (CS)
  - Expanding CS by searching for more options

Examples

- Consumer sequentially explores products + searches for more options
- Firm evaluates candidates + expands candidate pool by searching for more
- R&D: pursuing alternative technologies + searching for new ones to explore
- Researcher alternates between ongoing projects + searches for new ideas

Sequential experimentation with endogenous CS

- CS constructed gradually over time in response to information the DM collects
- Each period, DM either explores an alternative in CS or expands it
- Exploring alternative generates signal about its value (independent of other alternatives) and yields payoff
- Decision to expand CS (= search): costly and yields (stochastic) set of new alternatives as a function of state of the "search technology"
- Search technology may evolve over time based on past outcomes
    - e.g., state of search technology may be stationary (iid sets of new options)
    - or may evolve reflecting DM's beliefs about alternatives outside of CS

Results

- Characterization of optimal exploration and expansion policy
- Key properties of exploration/search dynamics: dependence on “search technology”
- Comparative statics

Applications

1. Clinical trials
2. Experimentation toward regulatory approval
3. Online consumer search (“Pandora’s boxes” w. endogenous set of boxes)

Results

- Characterization of optimal exploration and expansion policy
- Key properties of exploration/search dynamics: dependence on “search technology”
- Comparative statics

Applications

1. Clinical trials
2. Experimentation toward regulatory approval
3. Online consumer search (“Pandora’s boxes” w. endogenous set of boxes)

Literature

- CS formation: Eliaz Spiegler '11, Masatlioglu Nakajima Ozbay '12, Manzini Mariotti '14, Caplin Dean Leahy '18
- Sequential allocation of attention: Ke Shen Villas-Boas '16, Austen-Smith Martinelli '18, Ke Villas-Boas '19, Gossner, Steiner Stewart '19, Che Mierendorff '19, Liang Mu Syrgkanis '19
- Garfagnini Strulovici '16, Schneider Wolf '19, Fershtman Pavan '20
- Branching: Weiss '88, Weber '92, Keller Oldale '03
- Extensions of Pandora's boxes: Olszewski Weber '15, Choi Smith '16, Doval '18, Greminger '20

Outline

- Model
- Characterization, dynamics of exploration and expansion
- Applications
  1. Clinical trials
  2. Experimentation toward regulatory approval
  3. Online consumer search (+extension of Weitzman's '79 problem)

Model - decisions

- Discrete time: $t = 0, ..., \infty$
- Available alternatives in period t: $C_t = \{1, ..., n_t\}$ (C_0 exogenous)
- At each t, DM either
  1. Explores alternative in C_t
  2. Expands considerations set
  3. Opts-out: alternative $i = 0$ (fixed payoff equal to outside option)

Categories + learning

- Each alternative belongs to an observable category $\xi \in \Xi$
  - Characterizes alternative's experimentation technology and payoff process
  - Alternatives within same category are ex-ante identical
- Exploring alternative $\rightarrow$ learning about fixed unknown $\mu \in \mathbb{R}$, drawn from distr Γ_ξ
  - Observe a signal realization and update beliefs about μ
  - θ generic sequence of signal realizations

Exploration: states and payoffs

- "State" of an alternative: $\omega^P = (\xi, \theta) \in \Omega^P$
- $H_{\omega^P} \in \Delta(\Omega^P)$: distribution over Ω^P, given ω^P
- Payoff: $u(\omega^P)$
- Key assumptions:
  - Alternatives' state "frozen" unless DM explores them
  - Processes are independent of calendar time
  - Evolution of states independent across alternatives, conditional on category

Expansion: search technology

- Expansion of CS: costly + adds stochastic set of new alternatives
- State of *search technology*: $\omega^S = ((c_0, E_0), (c_1, E_1), ..., (c_m, E_m)) \in \Omega^S$
  - m: number of past searches
  - c_k: cost of k'th search
  - $E_k = (n_k(\xi) : \xi \in \Xi)$: result of k-th search

    $n_k(\xi)$: number of alternatives of category ξ discovered
- $H_{\omega^S} \in \Delta(\Omega^S)$: joint distribution over next (c, E), given ω^S
- Key assumptions:
  - Independence of calendar time
  - Search technology independent of θ (correlation though ξ)
- Stochasticity in search technology can capture
  - Learning about set of alternatives outside CS
  - Evolution of DM's ability to find new alternatives

State of decision problem + policies

- Period-t (overall) state: $\mathcal{S} \equiv (\omega^S, \mathcal{S}^P)$
  - ω^S: state of search technology
  - $\mathcal{S}^P : \Omega^P \to \mathbb{N}$ state of CS
    - $\mathcal{S}^P(\omega^P)$: number of alternatives in CS in state $\omega^P \in \Omega^P$
- *A policy* χ prescribes feasible decisions at all histories
- Policy χ is *optimal* if maximizes $\mathbb{E}^\chi \left[\sum_{t=0}^{\infty} \delta^t U_t | \mathcal{S}_0\right]$

Example: Clinical trials

- Exploring various medical treatments with unknown efficacy/safety
- DM sequentially chooses between treatments to administer
- Tradeoff - well-being of current patient vs value of learning about treatments
- Enrich this classic problem by endogenizing the DM's CS

Example: Clinical trials

- Each period ($t = 0, 1, ...$), physician chooses
  - which treatment to administer
  - or whether to search for additional treatments (to be added to the pool)
- Two categories of treatments: $\xi \in \Xi \equiv \{\alpha, \beta\}$
- Ex-ante, treatment from same category are identical
- Category-ξ treatments' efficacy $\mu^{\xi} \in \{0, 1\}$ unknown ex-ante, independent
- $p^{\xi}(\emptyset) = \Pr(\mu^{\xi} = 1)$ prior that a ξ-treatment is effective

Example: Clinical trials

- Outcome of treatment $s \in \{G, B\}$
- Using an effective ξ-treatment: $s = G$ w.p. $q^\xi \equiv \Pr(s = G|\mu^\xi = 1) \in (0, 1]$
- Using an ineffective ξ-treatment:: $s = B$ with certainty
- Given history $\theta = (s_1, s_2, ...)$, $p^\xi(\theta)$ posterior prob that the treatment is effective
- Payoff u from successful ξ-treatment: $v^\xi > 0$ if outcome is good, 0 otherwise
- Search for new treatment $\rightarrow$ identify ξ-treatment w.p. ρ^ξ, where $\rho^\alpha + \rho^\beta = 1$
- Cost of search: $c \geq 0$

Outline

- Model
- Characterization, dynamics of exploration and expansion
- Applications
  1. Clinical trials
  2. Experimentation toward regulatory approval
  3. Online consumer search (+extension of Weitzman's '79 problem)

Index of an alternative in CS (standard Gittins index):

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

- τ : stopping time (realization dependent)
- Interpretation: maximal expected discounted payoff, per unit of expected discounted time

Index for expansion of CS

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

- τ : stopping time
- π: choice among alternatives discovered after search launched and future searches

Index of an alternative in CS (standard Gittins index):

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

- τ : stopping time (realization dependent)
- Interpretation: maximal expected discounted payoff, per unit of expected discounted time

Index for expansion of CS

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

- τ : stopping time
- π: choice among alternatives discovered after search launched and future searches

Definition - Index policy χ^*

Expand CS at period t iff

$$\mathcal{I}_t^S(\omega^S) \ \geq \ \underbrace{\mathcal{I}_t^*(\mathcal{S}^P)}_{\substack{\text{maximal index among} \\ \text{available alternatives}}}$$

Otherwise, explore any alternative with index $\mathcal{I}_t^*(\mathcal{S}^P)$

Theorem 1 (optimal policy)

1. Optimal policy: index policy χ^* is optimal
2. Recursive structure: index of search can be written as

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

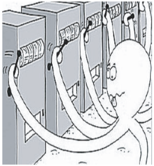
   1. τ^* is first time $s \geq 1$ at which $\mathcal{I}^S$ and all indices of alternatives brought in by search fall weakly below $\mathcal{I}^S(\omega^S)$
   2. expectations are wrt process induced by optimal policy χ^*
3. Value function: DM's expected (per-period) payoff under χ^* is

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

   - $\kappa(v)$ = minimal time, starting from initial state $\mathcal{S}_0$, till *all* indices $\leq v$

Methodology

- New proof of optimality of "index policies" for class of MAB problems where "arms" added as result of deliberate decision to search
- Related to "branching" lit: Weiss '88, Weber '92, Keller Oldale '03
- Key: proof yields recursive representation of index for expansion
  - + new representation of DM's payoff under optimal policy
- Central for deriving properties of dynamics, comparative statics, applications



Proof of Theorem 1: Road Map

1. Characterization of DM's payoff under index policy
2. Payoff function under index policy solves dynamic programming equation

Proof: Step 1

- $\kappa(v) \in \mathbb{N} \cup \{\infty\}$: minimal time until *all* indices drop weakly below $v \in \mathbb{R}_+$

Lemma 1

$$\underbrace{\mathcal{V}(\mathcal{S}_0)}_{\substack{\text{payoff under} \\ \text{index policy, starting} \\ \text{from state } \mathcal{S}_0}} = \int_0^\infty [1 - \underbrace{\mathbb{E}^{\chi^*}\left[\delta^{\kappa(v)} | \mathcal{S}_0\right]}_{\substack{\text{expected discounted} \\ \text{time till all indexes} \\ \text{drop weakly below } v}}] dv$$

-

Proof: Step 2

- $\mathcal{V}(\mathcal{S}_0)$ solves dynamic programming equation:

$$\mathcal{V}(\mathcal{S}_0) = \max\{\underbrace{V^S(\omega^S|\mathcal{S}_0)}_{\substack{\text{value from searching}\\ \text{and reverting}\\ \text{to index}\\ \text{policy thereafter}}}, \underbrace{\max_{\omega^P \in \{\hat{\omega}^P \in \Omega^P : \mathcal{S}_0^P(\hat{\omega}^P)>0\}} V^P(\omega^P|\mathcal{S}_0)}_{\substack{\text{value from exploring}\\ \text{alternative and}\\ \text{reverting to index}\\ \text{policy thereafter}}}\}$$

- Proof uses
  - representation of payoff under index policy from Lemma 1
  - decomposition of overall problem into collection of binary problems where choice is between single alternative (possibly search) and auxiliary fictitious alternative with fixed payoff

Implication for dynamics - I

1. Invariance of expansion to CS composition: at any period, expansion decision invariant in composition of CS, conditional on
   1. state ω^S of search technology
   2. value of highest index in current CS
2. IIA: at any period t, the choice between any pair of alternatives $i, j \in C_t$ is invariant in ω^S

Implication for dynamics - II

Definition:

A search technology is stationary if $(-c_k, E_k)$ drawn from fixed distribution, deteriorating if $(-c_k, E_k)$ is (FOSD) decreasing in k, and improving if $(-c_k, E_k)$ is (FOSD) increasing in k.

3. If search technology is stationary, for any two states $\mathcal{S}$, $\mathcal{S}^{'}$ at which DM expands CS, expected continuation payoff is the same
4. If search technology is stationary or improving and search is carried out at period t, DM never returns to any alternative in period-t CS
5. If search technology is stationary or deteriorating, decision to expand CS is the same as in a fictitious environment in which DM expects to have only one further opportunity to expand

- Model
- Characterization, dynamics of exploration and expansion
- Applications

  1. Clinical trials
  2. Experimentation toward regulatory approval
  3. Online consumer search (+extension of Weitzman's '79 problem)

Clinical trials

Optimal exploration/expansion policy

- Each treatment in CS assigned an index

$$\mathcal{I}^P(\xi,\theta) = \frac{\left(1-\delta+\delta q^{\xi}\right) p^{\xi}(\theta) q^{\xi} v^{\xi}}{1-\delta+\delta p^{\xi}(\theta) q^{\xi}}$$

- Expansion of treatment pool assigned index

$$\mathcal{I}^S = \frac{(1-\delta)\left(\sum_{\xi\in\{\alpha,\beta\}} \rho^{\xi}\mathbb{E}\left[\sum_{s=0}^{\tau^{\xi *}-1} \delta^s u_s|\xi\right] - c\right)}{1-\sum_{\xi\in\{\alpha,\beta\}} \rho^{\xi}\mathbb{E}\left[\delta^{\tau^{\xi *}}|\xi\right]}$$

($\tau^{\xi *}$ = first time that index of new ξ-treatment brought in by search $\leq \mathcal{I}^S$)

- Highest index determines decision at each period

Detrimental effect of improvement in a category

Consider an improvement in category α of treatments:

- $p^{\alpha}(\emptyset) \nearrow$, and/or $v^{\alpha} \nearrow$, and/or $q^{\alpha} \nearrow$

Improvement can lead to ex-ante *reduction* in expected discounted number of times α-treatments are administered.

- Improvement in α increases index $\mathcal{I}^P(\boldsymbol{\alpha}, \theta)$ of α-treatments, but also $\mathcal{I}^S$
  - Increase in $\mathcal{I}^P(\alpha, \theta)$ differs across histories of outcomes θ
  - $\mathcal{I}^S$ averages over histories at which a new α-category is administered
  - For some θ (e.g., after bad outcomes), increase in $\mathcal{I}^P(\alpha, \theta)$ may be *smaller* than increase in $\mathcal{I}^S$
- Search then shifts balance in CS in favor of β treatments (e.g., if $\rho^{\beta} > \rho^{\alpha}$)
- Can lead to an overall reduction in the usage of α-treatments

- Model
- Characterization, dynamics of exploration and expansion
- Applications
  1. Clinical trials
  2. Experimentation toward regulatory approval
  3. Online consumer search (+extension of Weitzman's '79 problem)

Experimentation toward regulatory approval

- Firm needs regulatory approval to sell its products
- Products differ in profitability to firm (known), and in their safety (unknown)
- Each product belongs to a category $\xi \in \Xi = \{\alpha, \beta\}$
- ξ-product is safe ($\mu^\xi = 1$) or not ($\mu^\xi = 0$)– ex-ante unknown to firm and regulator
  - Prior $p^\xi(\emptyset) = \Pr(\mu^\xi = 1)$
- Firm gets flow payoff $(1-\delta)v^\xi$ from selling approved ξ-product
- No value to firm in selling more than one product per period (e.g., substitutes)
- Each period, firm chooses between
  - experimenting with a product in its CS (at cost $\lambda^\xi(\theta)$)
  - expanding CS by searching for new products (at cost c)
  - selling an approved product

Experimentation/Expansion and approval

- Each experiment on ξ-product generates outcome $s \in \{G, B\}$
- Safe ξ-product: good outcome w.p. $q_1^\xi = \Pr(s = G|\mu^\xi = 1) \in (0, 1]$
- Unsafe ξ-product: bad outcome w.p $q_0^\xi = \Pr(s = B|\mu^\xi = 0)$, w. $q_1^\xi \geq 1 - q_0^\xi$
- Experimentation outcomes θ are public (Henry Ottaviani '19)
- $p^\xi(\theta) =$ posterior probability that a ξ-product is safe, given θ
- Expansion of CS yields single product, ρ^ξ prob that new product is of category ξ
- For each category ξ, product is approved iff $p^\xi(\theta) \geq \Psi^\xi \in (0, 1]$

Approval and firm's optimal policy

- Firm's goal: maximize expected discounted payoff from selling (approved) product, net of experimentation + search costs
- Firm's optimal policy: special case of the model, based on indices for experimentation and expansion
- Because experimenting with approved product is dominated by selling it, index of approved ξ-product is constant at $(1-\delta)v^{\xi}$
- Hence, approval of one of the firm's products ends its experimentation process
- What if regulator adopts policy relaxing approval standard for a category?

Changes in regulator's approval standard

Unintended effects of reducing a category's approval standard

Relaxation of category-α approval threshold can *reduce* the ex-ante prob that an α-product is approved

- Result hinges on endogeneity of the CS
- Relaxation of standard increases indices of α-products, but also index of search
- Index for search may increase more than the index of α-products that have yielded negative results
- Search then re-balances CS in favor of β-products, crowding out further evaluations of such α-products
- Can lead to reduction in ex-ante probability that α-products are approved

- Model
- Characterization, dynamics of exploration and expansion
- Applications
    1. Clinical trials
    2. Experimentation toward regulatory approval
    3. Online consumer search (+extension of Weitzman's '79 problem)

Weitzman '79 with endogenous set of boxes

- Category-ξ alternative characterized by (F^ξ, λ^ξ)
  - λ^ξ cost of opening box
  - F^ξ distr of box's value, v
- DM initially aware of only subset of alternatives - C_0
- Each period, DM either
  - expands CS
  - inspects an alternative to learn its value, or
  - stops and recalls prize v from inspected box, or takes outside option
- Expansion brings new box: ρ^ξ prob that search brings category-ξ box
- $c(m) =$ search cost, positive and increasing in # of past searches m
- Weitzman's "Pandora's boxes" problem: exogenous, fixed CS C_0 ($\rho^\xi \equiv 0$, $\forall \xi$)

Reservation price of a category-ξ box – defined as in Weitzman

$$\mathcal{I}^P(\omega^P) = \frac{-\lambda^\xi + \delta \int_{\frac{\mathcal{I}^P(\omega^P)}{1-\delta}}^{\infty} v \mathrm{d}F^\xi(v)}{1 + \frac{\delta}{1-\delta}\left(1 - F^\xi\left(\frac{\mathcal{I}^P(\omega^P)}{1-\delta}\right)\right)}$$

Reservation price of search/expansion

Define $\Xi(I) \equiv \{\xi \in \Xi : \mathcal{I}^P(\xi, \emptyset) > I\}$ (set of box categories w. reservation price $> I$).

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

- Optimal policy is based on comparison of independent reservation-prices (indices)
- Generalizes Weitzman's solution

Reservation price of a category-ξ box – defined as in Weitzman

$$\mathcal{I}^P(\omega^P) = \frac{-\lambda^\xi + \delta \int_{\frac{\mathcal{I}^P(\omega^P)}{1-\delta}}^{\infty} v \mathrm{d}F^\xi(v)}{1 + \frac{\delta}{1-\delta}\left(1 - F^\xi\left(\frac{\mathcal{I}^P(\omega^P)}{1-\delta}\right)\right)}$$

Reservation price of search/expansion

Define $\Xi(I) \equiv \left\{\xi \in \Xi : \ \mathcal{I}^P(\xi, \emptyset) > I\right\}$ (set of box categories w. reservation price $> I$).

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

- Optimal policy is based on comparison of independent reservation-prices (indices)
- Generalizes Weitzman's solution

Online consumer search

- Firms' ads listed in sequence, positions $m = 1, 2, \ldots$

1. Expansion $=$ reading ad displayed at next position
   - Each category $\xi \in \Xi$ corresponds to different firm
   - Reading next ad brings its product into CS
   - Reveals identity $\xi(m)$ of firm, drawn from stationary distr $\rho \in \Delta(\Xi)$
   - $c(m)$ cost of reading m'th result, $c(\cdot)$ non-decreasing
2. Opening box $=$ clicking to view product's page (learn v_m at cost $\lambda^{\xi(m)}$)
3. Stopping and choosing an opened box $=$ purchasing a product

- Optimal policy for consumer follows from the extension of Weitzman:
  - $\mathcal{I}^S(m)$: "reading index" (for decision to read m'th position)
  - $\mathcal{I}_m$: "clicking index" (for clicking m'th ad)
  - $(1-\delta)v_m$: "purchase index" (for purchasing product on m'th position)

Online search: Eventual purchase

- Choi Dai Kim '18 - static condition characterizing eventual purchase in Weitzman's setting (w. exogenously fixed CS)
  - Eventual purchase characterized by comparison of "effective values"
  - $w_m \equiv \min\{\mathcal{I}_m, v_m(1-\delta)\}$
  - Special case of our model where all products have already been read

- Define $d_m \equiv \min\{w_m, \mathcal{I}^{\mathcal{S}}(m)\}$ "discovery value"

- Outside option $\rightarrow$ position $m = 0$ (with $w_0 = d_0 = 0$)

Eventual purchase with endogenous CS

Consumer purchases product m if, for all $l \in \mathbb{N} \cup \{0\}$, $l \neq m$, $d_l < d_m$ (and only if $d_l \leq d_m$, for all $l \neq m$).

- Discovery values account for endogenous order in which various alternatives are read - can be used to study the effects of varying this order

Endogenizing click-through-rates (CTR)

- $CTR(m) \equiv \Pr(m\text{'s ad is clicked}|m\text{'s ad is read})$
- Important for sponsored search
- But connection between CTRs and positions typically exogenously assumed

Characterization of CTR

The CTR for each position $m \geq 1$ is given by

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

Adverse effects of additional ad space on firms' profits

- Three multi-product firms $\xi \in \Xi = \{A, B, C\}$
- Consumer's initial CS has three products, one from each firm $\xi = A, B, C$
- Searching online $\rightarrow$ consumer presented w. fourth ad, drawn from $\rho \in \Delta(\Xi)$
  - i.e., fourth ad belongs to one of the three firms (realized firm's 2nd product)

Additional ad space may reduce firm's profits

An increase in the probability ρ^{ξ} that search brings an additional firm-ξ product may reduce firm ξ's ex-ante expected profits.

- Increase in prob search brings additional firm-ξ product may reduce index of search
- Can induce consumer to click on firm ξ's competitors before searching
- Reduces prob that one of firm ξ's product is selected, and hence its profits

Conclusion

- Study sequential learning with endogenous set of alternatives
- CS constructed gradually in response to arrival of info (dynamic micro-foundation)
- Key tradeoff: exploring alternatives already in CS vs expanding CS
- Characterize optimal policy, implications for dynamics, comparative statics
- Useful for applications where DM unaware of all feasible options from the start
  - limited attention
  - sequential provision of information by another party
- Applications: clinical trials, persuading a regulator, consumer search, recruitment
- Special case: extension of Weitzman '79 Pandora's boxes problem

THANKS! ☺

THANKS!

☺

Meta Arms

- Arm 1:
  - 1,000 first time
  - $\lambda \in \{1, 10\}$ subsequent times (equal probability, perfectly persistent)
- Arm 2 (Meta Arm) can be used in two modes
  - 2(A): 100 first time, 0 thereafter
  - 2(B): 11 each period
- Selection of Arm 2's mode is irreversible
- Optimal policy ($\delta = .9$):
  - start w. Arm 1, and then
    - If $\lambda = 10$, use arm 2 in mode 2(A) for one period, followed by arm 1 thereafter
    - If $\lambda = 1$, use arm 2 in mode 2(B) thereafter
- No index representation, regardless of how we define the "index"

Go back

Interpretation of reservation prices

Suppose only two alternatives:

- Alternative i characterized by ξ (which determines (F^{ξ}, λ^{ξ})), and
- hypothetical alternative, j, with known value v_j

Reservation price of box i is value v_j for which DM is indifferent between

- taking j right away
- inspecting i while maintaining option to recall j once v_i is discovered

Interpretation of reservation price of search $\mathcal{I}^S$

Suppose only two options:

- hypothetical alternative, j, with known value v_j
- option of expanding the CS

Reservation price of search is value v_j for which DM is indifferent between

- taking j right away,
- expanding the CS, maintaining the option to take j either
  - once ξ of new alternative is discovered and $v_j \geq \mathcal{I}^P(\xi, \emptyset)$
  - or if $v_j < \mathcal{I}^P(\xi, \emptyset)$, after value v_i of new alternative is learned and $v_i \leq v_j$

Go back

Policy: formal definition

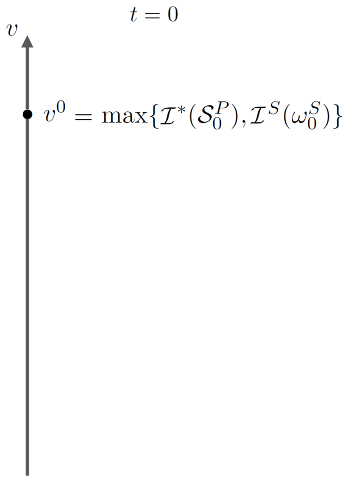
- Period-t decision: $d_t \equiv (x_t, y_t)$
  - $x_{it} = 1$ if alternative i explored; $x_{it} = 0$ otherwise
  - $y_t = 1$ if search; $y_t = 0$ otherwise
  - Sequence of decisions $d = (d_t)_{t=0}^{\infty}$ *feasible* if, for all $t \geq 0$:
    - $x_{jt} = 1$ only if $j \in I_t$
    - $\sum_{j \in I_t} x_{jt} + y_t = 1$
- Rule χ governing feasible decisions $(d_t)_{t \geq 0}$ is a policy iff sequence of decisions $\{d_t^{\chi}\}_{t \geq 0}$ under χ is $\{\mathcal{F}_t^{\chi}\}_{t \geq 0}$-adapted, where $\{\mathcal{F}_t^{\chi}\}_{t \geq 0}$ is natural filtration induced by χ

Go back

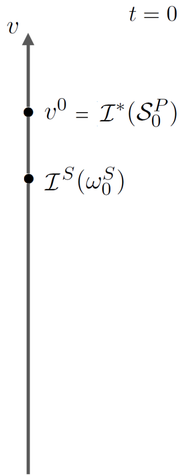
Proof of Lemma 1

- $v^0 = \max\{\mathcal{I}^*(\mathcal{S}_0^P), \mathcal{I}^S(\omega_0^S)\}$
- t^0: first time all indices (including search) strictly below v^0 ($t^0 = \infty$ if event never occurs)
- $\eta(v^0)$: discounted sum of payoffs, net of search costs, till t^0 (includes payoffs from newly added alternatives)
- $v^1 = \max\{\mathcal{I}^*(\mathcal{S}_{t^0}^P), \mathcal{I}^S(\omega_{t^0}^S)\}$ (note: $t^0 = \kappa(v^1)$)
- ...
- $\eta(v^i)$: net payoff between $\kappa(v^i)$ and $\kappa(v^{i+1}) - 1$
- Stochastic sequence of values $(v^i)_{i \geq 0}$, times $(\kappa(v^i))_{i \geq 0}$, and discounted net payoff $(\eta(v^i))_{i \geq 0}$

Proof of Lemma 1



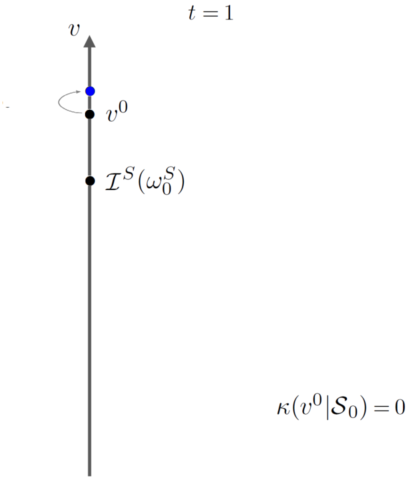
Proof of Lemma 1



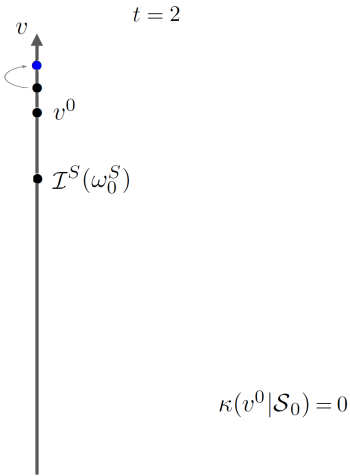
$v^0 = \text{index of arm}$

$\kappa(v^0|\mathcal{S}_0) = 0$

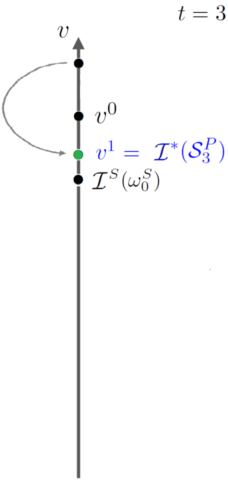
Proof of Lemma 1



Proof of Lemma 1



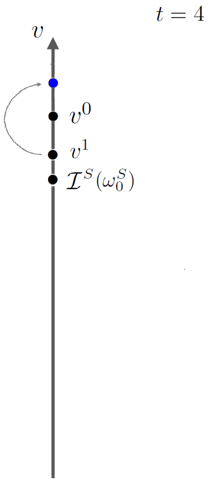
Proof of Lemma 1



$\kappa(v^0|\mathcal{S}_0) = 0$

$t^0 = \kappa(v^1|\mathcal{S}_0) = 3$

Proof of Lemma 1

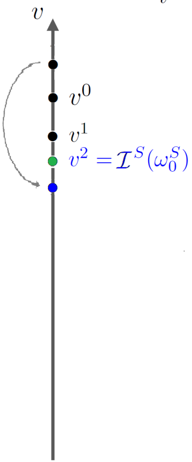


$\kappa(v^0|\mathcal{S}_0) = 0$

$t^0 = \kappa(v^1|\mathcal{S}_0) = 3$

Proof of Lemma 1

$t = 5$

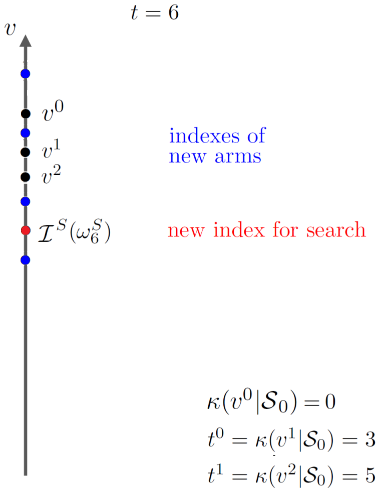


$\kappa(v^0|\mathcal{S}_0) = 0$

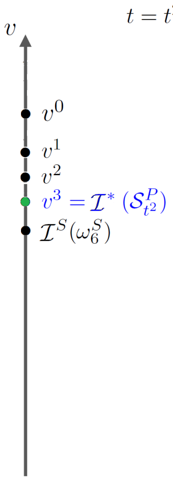
$t^0 = \kappa(v^1|\mathcal{S}_0) = 3$

$t^1 = \kappa(v^2|\mathcal{S}_0) = 5$

Proof of Lemma 1



Proof of Lemma 1



$\kappa(v^0|\mathcal{S}_0) = 0$

$t^0 = \kappa(v^1|\mathcal{S}_0) = 3$

$t^1 = \kappa(v^2|\mathcal{S}_0) = 5$

$t^2 = \kappa(v^3|\mathcal{S}_0)$

Proof of Lemma 1

- (Average) payoff under index policy:

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

- Starting at $\kappa(v^i)$, optimal stopping time in index defining v^i is $\kappa(v^{i+1})$
  - if v^i is index of alternative, $\kappa(v^{i+1})$ is first time its index drops below v^i
  - if v^i is index of expansion, $\kappa(v^{i+1})$ is first time search index + index of all alternatives discovered after $\kappa(v^i)$ drop below v^i

- Hence, v^i = expected discounted sum of net payoffs, per unit of expected discounted time, from $\kappa(v^i)$ until $\kappa(v^{i+1}) - 1$:

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

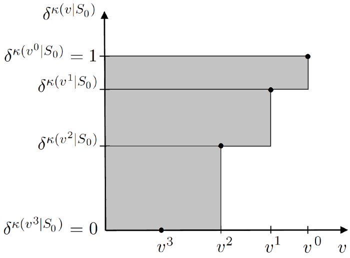
- Same true if multiple alternatives and/or search have index equal to v^i at $\kappa(v^i)$

Proof of Lemma 1

- Plugging in expression for v^i,

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

$$\sum_{i=0}^{\infty} v^i \left(\delta^{\kappa(v^i|\mathcal{S}_0)} - \delta^{\kappa(v^{i+1}|\mathcal{S}_0)}\right)$$



- Therefore,

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

Go back

Proof of DP

- Want to show that $\mathcal{V}(\mathcal{S}_0)$ solves dynamic programming equation:

$$\mathcal{V}(\mathcal{S}_0) = \max\{ \underbrace{V^S(\omega^S|\mathcal{S}_0)}_{\substack{\text{value from searching} \\ \text{and reverting} \\ \text{to index} \\ \text{policy thereafter}}}, \underbrace{\max_{\omega^P \in \{\hat{\omega}^P \in \Omega^P : \mathcal{S}_0^P(\hat{\omega}^P) > 0\}} V^P(\omega^P|\mathcal{S}_0)}_{\substack{\text{value from pulling} \\ \text{physical arm and} \\ \text{reverting to index} \\ \text{policy thereafter}}} \}$$

Auxiliary alternatives

- $e(\omega_M^A)$: state with single auxiliary alternative yielding fixed payoff M
- Note: $\kappa(v|\underbrace{\mathcal{S}_0 \vee e(\omega_M^A)}_{\mathcal{S}_0\ +\ \text{auxiliary arm}}) = \begin{cases} \kappa(v|\mathcal{S}_0) & \text{if } v \geq M \\ \infty & \text{otherwise} \end{cases}$
- From Lemma 1, payoff from index policy when auxiliary alternative added:

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

Auxiliary alternatives

$$\underbrace{D^S(\omega^S|e(\omega^S)\vee e(\omega_M^A))}_{\substack{\text{loss from starting}\\ \text{with search given only}\\ \text{search + auxiliary}\\ \text{arm}}} \equiv \underbrace{\mathcal{V}(e(\omega^S)\vee e(\omega_M^A))}_{\substack{\text{value under index}\\ \text{policy given only}\\ \text{search + auxiliary}\\ \text{arm}}} - \underbrace{V^S(\omega^S|e(\omega^S)\vee e(\omega_M^A))}_{\substack{\text{value of searching}\\ \text{and reverting to index}\\ \text{policy given only search}\\ \text{+ auxiliary arm}}}$$

$$= \begin{cases} 0 & \text{if } M \leq \mathcal{I}^S(\omega^S) \\ > 0 & \text{if } M > \mathcal{I}^S(\omega^S) \end{cases}$$

Similarly, for physical alternative in state ω^P:

$$D^P(\omega^P|e(\omega^P)\vee e(\omega_M^A)) = \begin{cases} 0 & \text{if } M \leq \mathcal{I}^P(\omega^P) \\ > 0 & \text{if } M > \mathcal{I}^P(\omega^P) \end{cases}$$

Proof that $\mathcal{V}$ solves Bellman eq

- Can show ("tedious"): $D^S(\omega^S|\boldsymbol{\mathcal{S}_0}) = \int_0^{\mathbf{v^0}} D^S(\omega^S|\mathbf{e(\boldsymbol{\omega}^S)} \vee \mathbf{e(\boldsymbol{\omega}^A_M)})\mathrm{d}\mathbb{E}\delta^{\kappa(M|\mathcal{S}_0^P)}$
- Hence: $D^S(\omega^S|\mathcal{S}_0) = 0$

$$\iff D^S(\omega^S|e(\omega^S) \vee e(\omega^A_M)) = 0,\ \forall M \in [0, \max\{\mathcal{I}^*(\mathcal{S}_0^P), \mathcal{I}^S(\omega^S)\}]$$
$$\iff \mathcal{I}^*(\mathcal{S}_0^P) \leq \mathcal{I}^S(\omega^S)$$

loss from starting with search $= 0$ iff search has largest index, and > 0 otherwise

- Similarly, $D^P(\omega^P|\mathcal{S}_0) = 0 \iff \mathcal{I}^P(\omega^P) = \mathcal{I}^*(\mathcal{S}_0^P) \geq \mathcal{I}^S(\omega^S)$
- Hence, $\mathcal{V}(\mathcal{S}_0) = \max\left\{V^S(\omega^S|\mathcal{S}_0), \max_{\omega^P \in \{\hat{\omega}^P \in \Omega^P : \mathcal{S}_0^P(\hat{\omega}^P) > 0\}} V^P(\omega^P|\mathcal{S}_0)\right\}$
- $\mathcal{V}(\mathcal{S}_0)$ solves dynamic programming equation (hence index policy optimal) ■

Validation

- Assumption: For any $\mathcal{S}$, and policy χ,

$$\lim_{t \to \infty} \delta^t \mathbb{E}^\chi \left[\sum_{s=t}^{\infty} \delta^s \left(\sum_{j=1}^{\infty} U_s \right) | \mathcal{S} \right] = 0$$

  - Solution to DP equation coincides with value function
  - Assumption satisfied if payoffs uniformly bounded
  - Also compatible with unbounded payoffs. E.g., alternatives are sampling processes, with payoffs drawn from Normal distribution with unknown mean

Go back