# Searching for Arms: Experimentation with Endogenous Consideration Sets

Daniel Fershtman, Alessandro Pavan

# Experimentation with endogenous set of alternatives

- Settings where decision maker (DM) alternates between
  - exploring alternatives in consideration set (CS)
  - expanding CS by searching for new alternatives
- Examples
  - **Bidders' solicitation in procurement auctions**
    - buyer alternates between experimenting with incumbents and searching for new suppliers
  - **Consumer search**
    - consumer sequentially explores products + searches for new ones
  - **Recruitment**
    - recruiter evaluates candidates + expands candidate pool

# This paper

- General multi-armed-bandit (MAB) model
  - set of "arms" expanded endogenously over time
  - exploration of alternatives
    - generates info about their value
    - yields payoffs
  - expansion of CS (= search)
    - costly
    - brings stochastic set of new alternatives
    - search outcome determined by past searches

# Results

- Optimal policy: Index solution
  - recursive characterization of index for CS expansion
- Dynamics of
  - CS expansion
  - experimentation
- Application: **bidder solicitation in procurement auctions**

Other applications

# 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
- **Branching** in MAB: Weiss '88, Weber '92, Keller Oldale '03
- **Solicitation in auctions**: Lauermann Wolinsky '25

# Plan

- General MAB model with endogenous CS
- Application: bidders' solicitation in auctions

# Model

# Decisions

- Discrete time: $t = 0, ..., \infty$
- Period-$t$ CS, $C_t = \{1, ..., n_t\}$: available alternatives
  - $C_0$ exogenous
- At each $t$, DM
  1. either explores one alternative in $C_t$
  2. or expands CS (i.e., searches for new alternatives)
  3. or opts-out: alternative $i = 0$ (fixed payoff equal to outside option)

# Categories

- Each alternative belongs to category $\xi \in \Xi$ encoding
  - signal process
  - payoff process
- Alternatives from same category: ex-ante identical
- Convenient way of modeling "search technology"

# Exploration

- Exploration of alternative of category $\xi$:
  - signal about unknown parameter $\mu \in \mathbb{R}$ (drawn from $\Gamma_\xi$)
  - $\mu$: encodes information about future signals and payoffs
- $\theta = (\vartheta_1, ..., \vartheta_m)$: generic sequence of signal realizations
- State of an alternative:

$$\omega^P = (\xi, \theta) \in \Omega^P$$

  - $\xi$: revealed upon arrival
  - $\theta$: generated by subsequent explorations

# Exploration

- $H_{\omega^P} \in \Delta(\Omega^P)$: distribution over next state, given current state
- (Expected) flow payoff: $u(\omega^P)$
- Key assumptions:
  - state $\omega^P$ changes only when alternative is explored
  - processes time-homogeneous (independent of calendar time)
  - states evolve independently across alternatives

# CS expansion: search technology

- CS Expansion (search): brings 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 $\xi$ discovered

# CS expansion: search technology

- $H_{\omega^S} \in \Delta(\Omega^S)$: distribution over next state of search technology, given current state
- Key assumptions
  - independence of calendar time
  - independence of alternatives' idiosyncratic signals $\theta$
- Serial correlation in search technology
  - learning about alternatives outside CS
  - evolution of DM's ability to find new alternatives

# Overall State and DM's Problem

- Period-$t$ overall state: $\mathcal{S} \equiv (\omega^S, \mathcal{S}^P)$
  - $\omega^S$: state of search technology
  - $\mathcal{S}^P : \Omega^P \to \mathbb{N}$: state of current CS
  - $\mathcal{S}^P(\omega^P)$: number of alternatives in CS in state $\omega^P \in \Omega^P$

# DM's Problem

- Choose policy $\chi$ prescribing for all histories
  - expand CS
  - which alternative in CS to explore

- Policy $\chi$ **optimal** if it maximizes

$$\mathbb{E}^{\chi}\left[\sum_{t=0}^{\infty}\delta^{t}U_{t}|\mathcal{S}_{0}\right]$$

  where $U_t$ are flow payoffs (net of all costs)

# Optimal policy

# Gittins indexes (for alternatives)

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

- $\tau$**: stopping time** (realization dependent)

- Interpretation: maximal expected discounted payoff, per unit of expected discounted time

# CS Expansion Index

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

- $\tau$: stopping time
- $\pi$: choice over exploration of alternatives discovered via current and future searches and future CS expansions

# Index Policy

- At each period $t$, index policy $\chi^*$
  - expand CS iff

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

  - explore one of alternatives with highest index $\mathcal{I}_t^*(\mathcal{S}^P)$ otherwise

# Index Policy

**Theorem**

- *Index policy* $\chi^*$ *optimal*
- *Index of search/CS expansion has recursive structure*

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

- $\tau^*$: *first time* $s \geq 1$ *at which* $\mathcal{I}^S$ *and all indices of alternatives brought to CS by current and future searches fall weakly below* $\mathcal{I}^S(\omega^S)$

- *expectations under process induced by index policy* $\chi^*$

# Proof of Theorem 1: Road Map

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

# Implications for dynamics

# Dynamics under optimal policy

1. Expansion decision depends only on
   - state $\omega^S$ of search technology
   - value of highest index among alternatives in CS
2. Independence of Irrelevant Alternatives: choice between any pair of alternatives $i, j \in C_t$ invariant in $\omega^S$

# Dynamics under optimal policy

**Definition**

Search technology

stationary: $(-c_k, E_k)$ drawn from fixed distribution

deteriorating: $(-c_k, E_k)$ drawn from distribution (FOSD) decreasing in $k$

improving: $(-c_k, E_k)$ drawn from distribution (FOSD) increasing in $k$

1. Stationary technology: for any $\mathcal{S}$, $\mathcal{S}'$ at which DM expands CS, expected continuation payoff same
2. Stationary or improving technology: if search carried out at $t$, DM never returns to any alternative in period-$t$ CS
3. Stationary or deteriorating technology: decision to expand CS same as in fictitious environment with only one further expansion

# Meta Arms

# "Meta Arms"

Can we generalize this further?

- treat "arms" as "meta-arms," each corresponding to own decision problem
- solve decision problem corresponding to each meta-arm separately
- assign index to each meta arm
- each period select meta-arm with highest index
- Typically suboptimal...

# "Meta Arms"

Arm 1:

- 1,000 first time
- $\lambda \in \{1, 10\}$ each subsequent time (equal probability, perfectly persistent)

Arm 2: Meta Arm that can be used in two modes:

- 2(A): 100 first time, 0 thereafter
- 2(B): 11 each period
- Selection of Arm 2's mode irreversible.
- Arm 2 has its own, independent decision problem

# "Meta Arms"

Optimal policy ($\delta = 0.9$)

- Start w. Arm 1,
  - if $\lambda = 10$, Arm 2 in mode 2(A) for one period, then Arm 1 thereafter
  - if $\lambda = 1$, Arm 2 in mode 2(B) thereafter
- No index representation, regardless of how "indexes" defined

# Application: bidders' solicitation

# Procurement problem

- Buyer repeatedly procures from multiple suppliers
- Set of suppliers expanded through
  - advertising
  - consultant soliciting bids
- Buyer faces uncertainty over
  - set of potential suppliers (how many firms are in mkt)
  - quality of firms' services
  - firms' cost of supplying goods/services
- Procurement design problem $\rightarrow$ naturally maps into MAB model with endogenous CS

# Uncertainty over pool of possible suppliers

- Total number of firms in mkt: $N \in \mathbb{N}$
- Buyer initially believes $N$ drawn from cdf $F$
  - $f(N)$: prob $N$ firms
- Each new search
  - identifies each firm in mkt not identified yet with prob $\rho \in (0, 1)$
  - given number of firms discovered, $n$, buyer updates about remaining number of firms $N - n$ in mkt
- Search cost: procurement paused for one period

# Ultra-log concave/convex distributions

**Definition**

Distribution $F$ *UL-concave* (*UL-convex*) if

$$L(N) \equiv \frac{(N+1)f(N+1)}{f(N)}$$

decreasing (increasing) in $N$

- $F \sim \text{Poisson}(\lambda)$
  - $L(N) = \lambda$
  - posterior over remaining firms in mkt: $\text{Poisson}(\lambda(1-\rho)^m)$
  - decreasing in number of past searches, $m$
  - invariant in number of firms discovered, $n$
- UL-concave (UL-convex): more (less) measure around mode than Poisson with same mean

# Firms' costs

- Each firm incurs cost $c$ each time it supplies
- $c$ drawn from cdf $G$
  - independently across firms (and all other r.v.)
  - constant over time
  - privately observed by firm
- $G/g$ increasing
- "virtual cost":

$$\gamma(c) \equiv c + G(c)/g(c)$$

# Firms' service quality

- Quality of each firm's service $\mu \in \{\underline{\mu}, \overline{\mu}\}$
  - $\Pr(\mu = \overline{\mu}) = p$
  - independent across firms (and all other r.v.)
- Neither buyer nor firm know $\mu$
- Buyer gradually learns $\mu$ through experimentation
- **No-news-bad-news**:
  - each time buyer procures, signal $\vartheta \in \{0, 1\}$
  - $\Pr(\vartheta = 0 | \mu = \underline{\mu}) = 1$
  - $\Pr(\vartheta = 1 | \mu = \overline{\mu}) = q \in (0, 1)$
- $\vartheta$ also buyer's flow payoff
  - expected flow payoff, given $\theta = (\vartheta_0, ..., \vartheta_m)$: $u(\theta) = q\Pr(\mu = \overline{\mu}|\theta)$

# Procurement Mechanisms

- Newly identified firms invited to join and submit bid
- Bid determines
  - payment from buyer each time firm supplies
  - lump-sum transfer $T$ (at joining)
- Firms identified in period $t$ submit bids at end of period $t$
- Buyer maximizes
  - welfare (expected discounted value of services net of firms' costs)
  - profits (expected discounted value of services net of payments)

# Optimal Mechanisms: Index Policy

**Lemma**

*Under both welfare- and profit-maximizing mechanism, buyer's decisions governed by* **index policy**.

- State of each firm: $\omega^P(b, \theta)$
  - category $\xi = b$: endogenous and equal to bid
- Flow cost of procuring from firm with bid $b$
  - welfare maximization: $b$
  - profit maximization: $\gamma(b) \equiv b + G(b)/g(b)$
- Optimal lump-sum transfers $T$
  - truthful bidding ($b = c$)
  - firms informed of state $\mathcal{S}$ of mechanism prior to bidding

# Search index

**Lemma**

*Under both welfare and profit maximization, $\mathcal{I}^S$*

1. *decreasing in number of past searches, $m$*
2. *decreasing (increasing) in number of firms identified, $n$, if $F$ UL-concave (UL-convex)*

- *Posterior over $N - n$*
  - *less optimistic (in MLR–order) as $m$ increases*
  - *UL-concave: more pessimistic as $n$ increases*
  - *UL-convex: more optimistic as $n$ increases*

# Positive implications I

**Proposition**

*Under both welfare and profit maximization:*

1. *Poisson or UL-concave $F$: buyer becomes more lenient as $m$ increases*
2. *UL-convex $F$: buyer becomes less lenient as $n$ increases*

- No-news-bad-news
  - leniency: # of failed attempts $A(b; m, n)$ granted before next expansion
- More generally
  - deterioration in index $\mathcal{I}^P$ tolerated before next expansion

# Positive implications II

**Proposition**

*Firms joining after larger m earn higher expected profits, fixing*

- *number of competitors n at time of joining*
- *firms' costs*
- *history of experimentation outcomes*

- The longer it took the buyer to assemble supplier pool, the higher a firm's expected profits when joining

# Normative implications

- $A_W(c; m, n)$: # of failed attempts under welfare max
- $A_{PR}(c; m, n)$: # of failed attempts under profit max

$$\Delta(c; m, n) \equiv A_W(c; m, n) - A_{PR}(c; m, n)$$

- Let $\hat{c}(m, n)$ solve

$$\mathcal{I}_W^S(m, n) - \mathcal{I}_{PR}^S(m, n) = G(c)/g(c)$$

# Normative implications

**Lemma**

*Following true:*

$$\Delta(c; m, n) \cdot (c - \hat{c}(m, n)) > 0$$

- Prior to $(m+1)$-th expansion
  - low-cost firms: inefficiently **large** number of failed attempts
  - high-cost firms: inefficiently **small** number of failed attempts
- $\hat{c}(m, n)$: non-monotone in $m$

# Normative implications

**Proposition**

*Following true:*

1. *Low-cost firms* ($c < \hat{c}(m, n)$) *contribute to inefficient entrenchment— buyer delays solicitation of new suppliers*

2. *High-cost firms* ($c > \hat{c}(m, n)$) *contribute to inefficient impatience with incumbents, prompting premature search*

- Analyst knowing all bids below $\hat{c}(m, n)$ predicts excessive delay in expansion of supplier pool

# Normative implications

- Contrary to standard screening
  - **all firms contribute to inefficiencies**
  - low-cost firms: inefficient delay in expansion of supplier pool
  - high-cost firms: excessive early expansions

# Conclusions

# Conclusions

- Experimentation
  - set of alternatives constructed in response to past search/experimentation outcomes
- Optimality of index policy
  - recursive CS expansion index
- Properties of
  - CS expansion dynamics
  - experimentation dynamics
- Applications
  - bidders' solicitation
  - online consumer search
  - recruitment
  - ...

# THANKS!

# Other applications

- **Online Consumer Search**
  - generalization of Weitzman "Pandora's boxes" problem to endogenous set of boxes
  - special case of MAB with immediate resolution of uncertainty
    - endogenous click-through-rates
    - foundation for clicking behavior observed in data
    - implications for sponsored-search auctions

# Other applications

- **Soft Affirmative Action and Minority Recruitment**
  - endogenous candidate pool
  - policies aimed at helping minorities (affirmative actions)
    - work with exogenous CS
    - can hurt minorities with endogenous CSE

Go back

# Proof: Step 1

**Lemma**

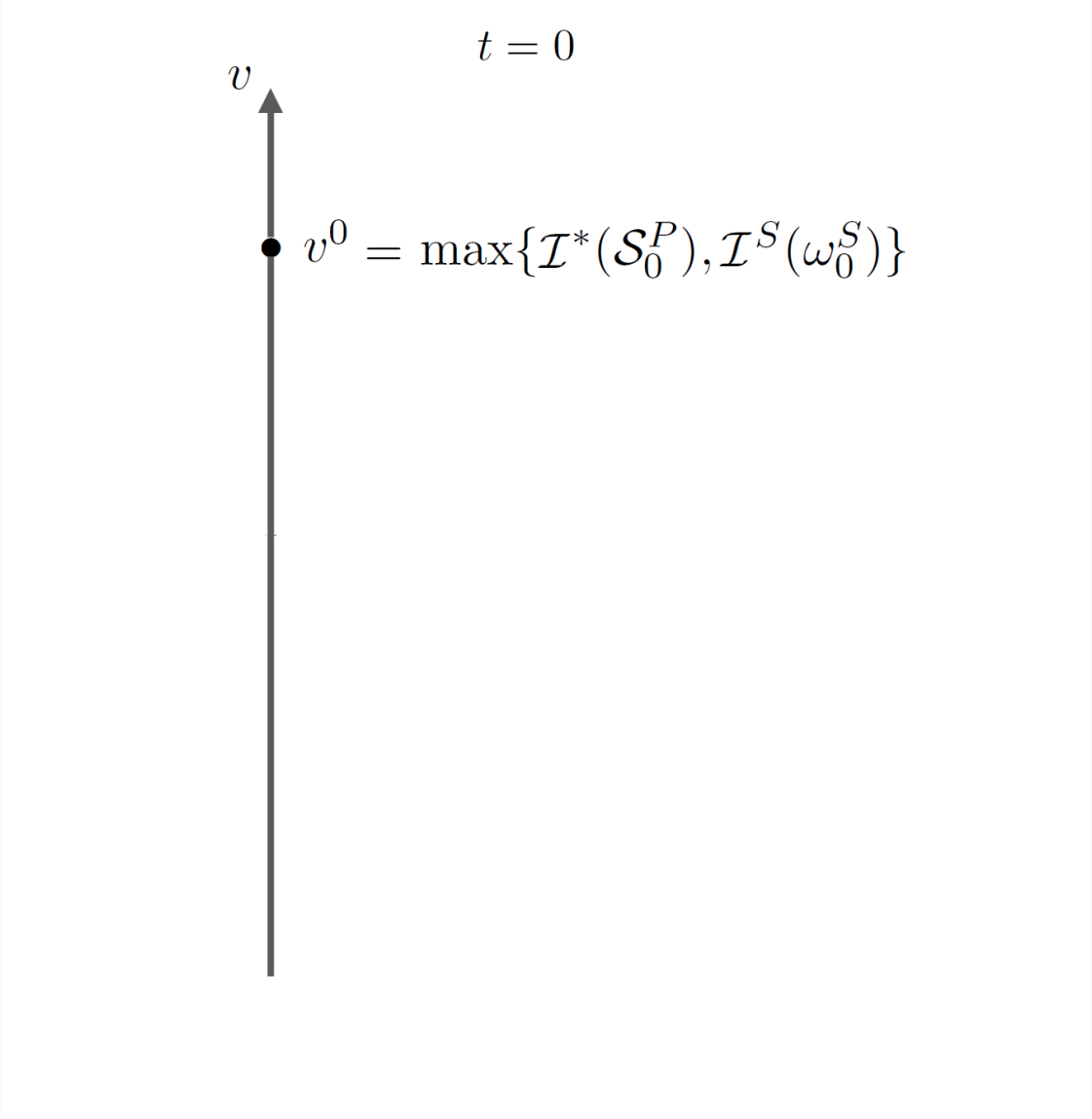
*Let $\kappa(v) \in \mathbb{N} \cup \{\infty\}$ be minimal time until all indices fall weakly below $v$. Then:*

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

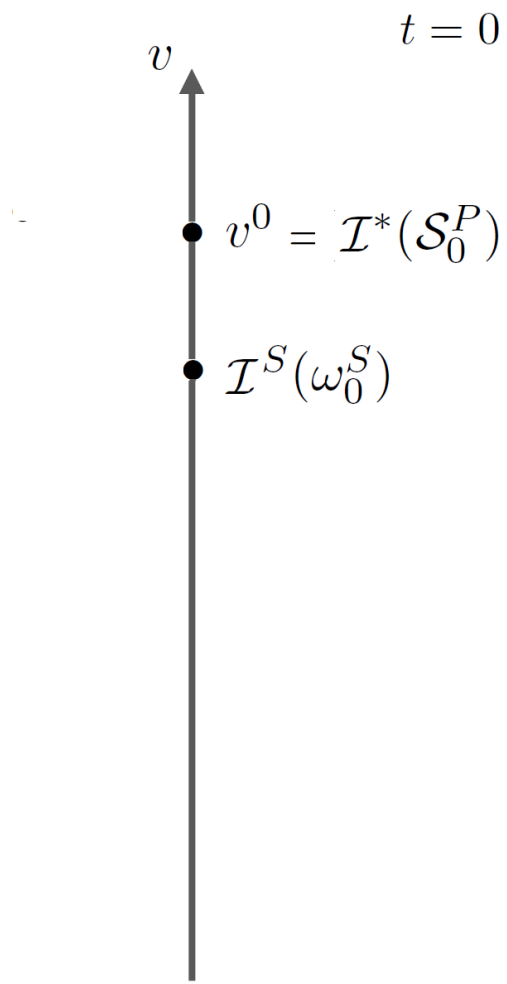
# Proof of Lemma 1

- $v^0 = \max\{\mathcal{I}^*(\mathcal{S}_0^P), \mathcal{I}^S(\omega_0^S)\}$ .
- $t^0$: 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



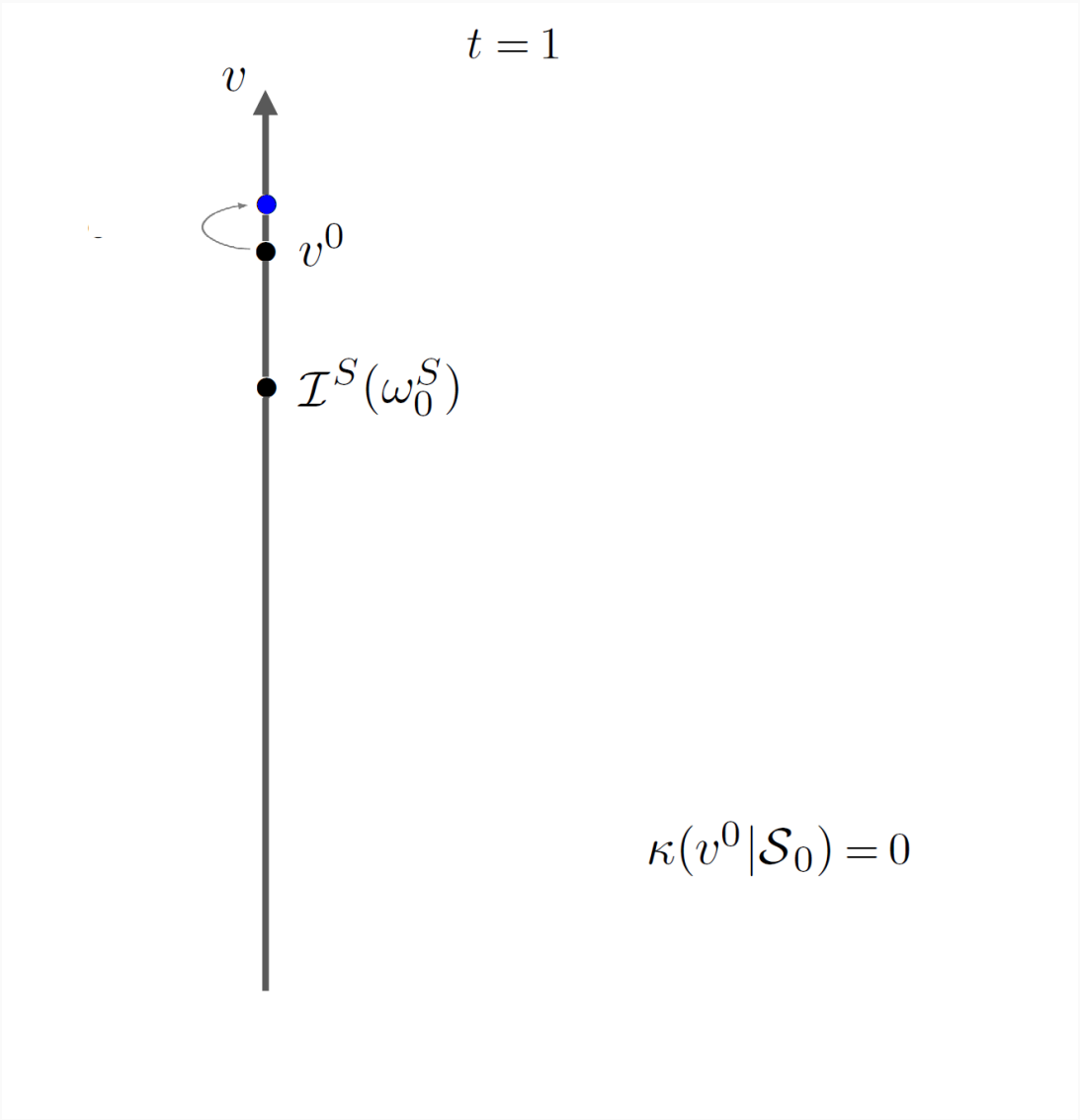
# Proof of Lemma 1



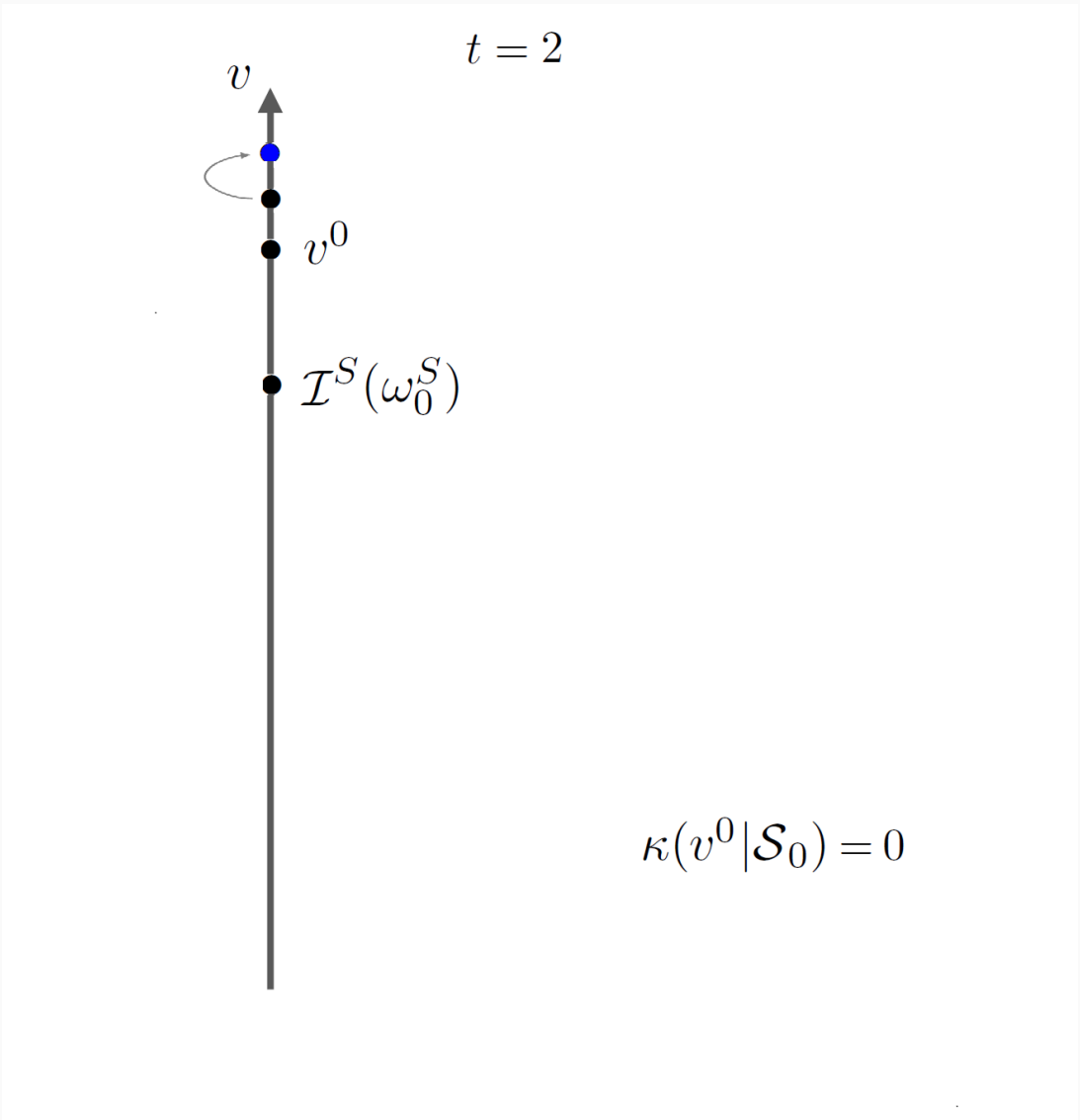

$v^0 = \text{index of arm}$

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

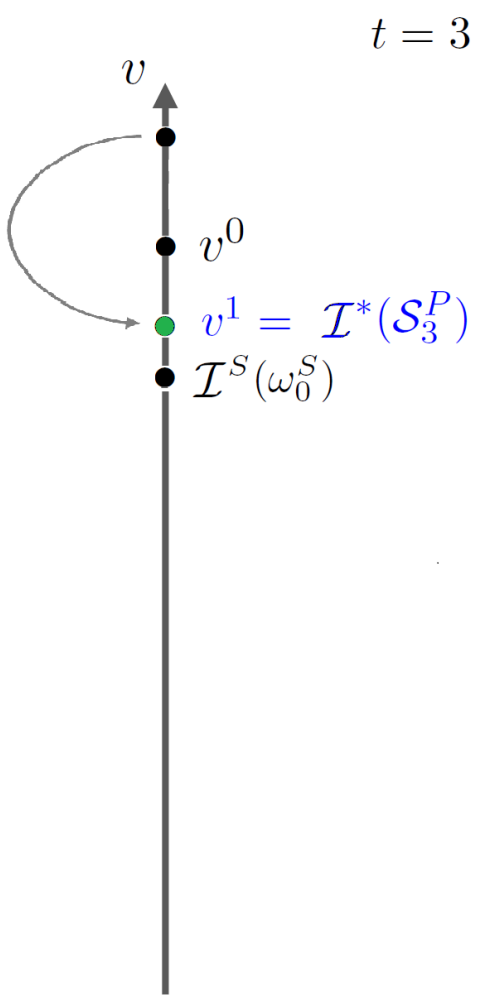
# Proof of Lemma 1



# Proof of Lemma 1



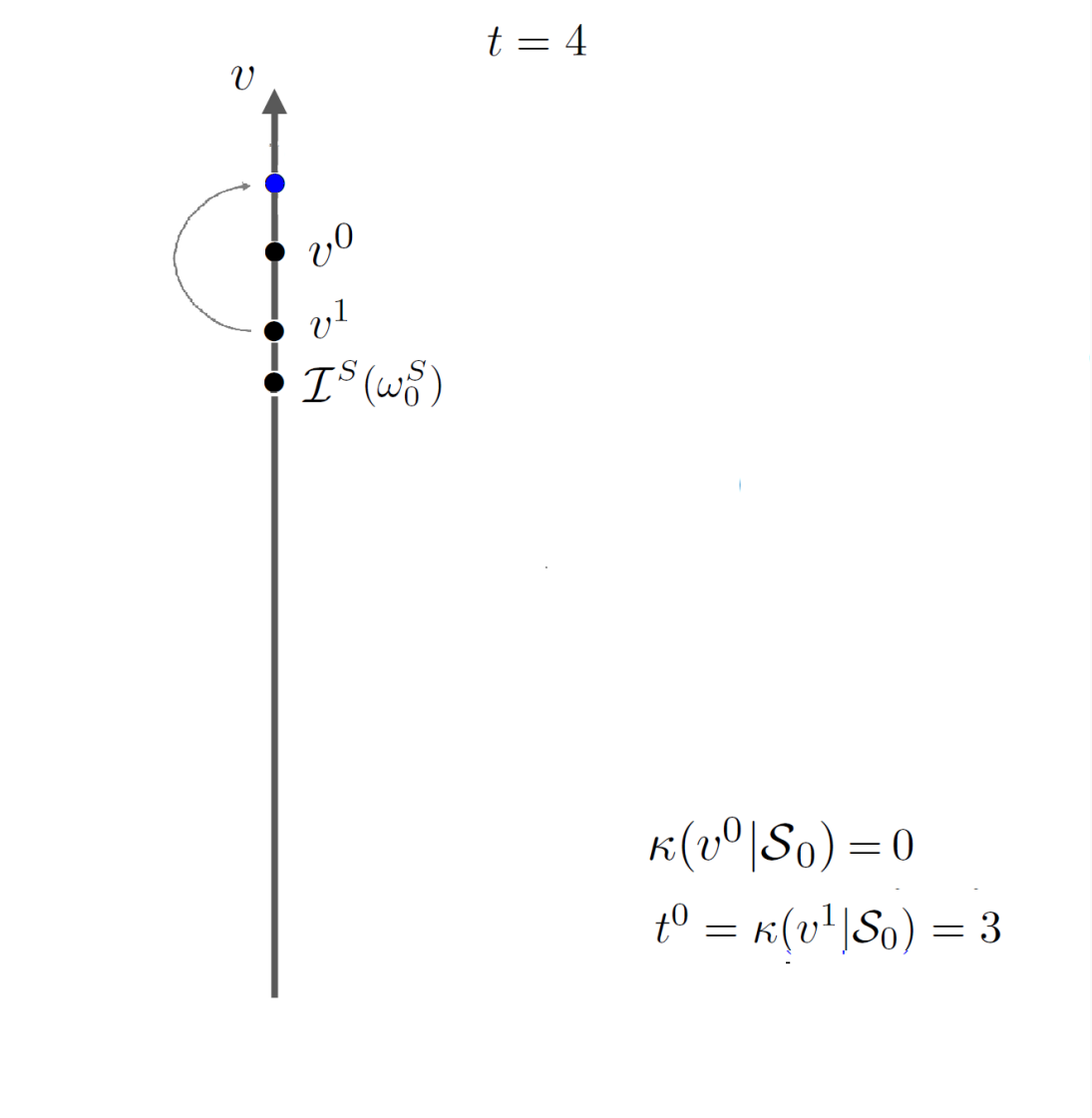
# Proof of Lemma 1



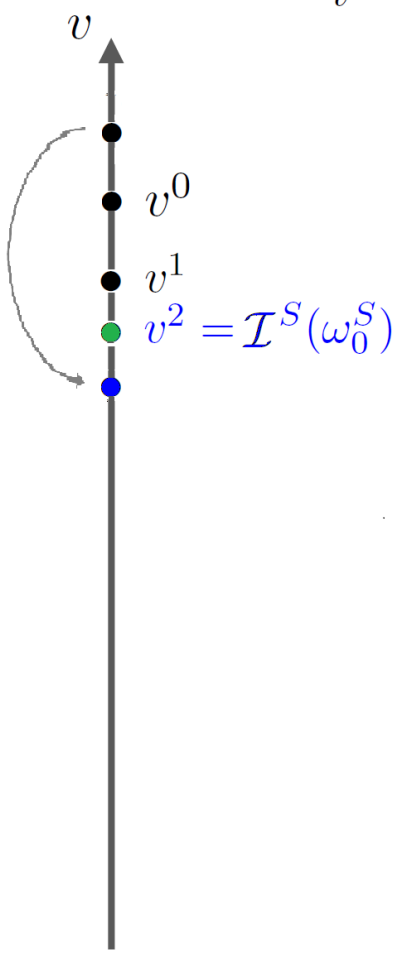

$$\kappa(v^0|\mathcal{S}_0) = 0$$

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

# Proof of Lemma 1



# Proof of Lemma 1

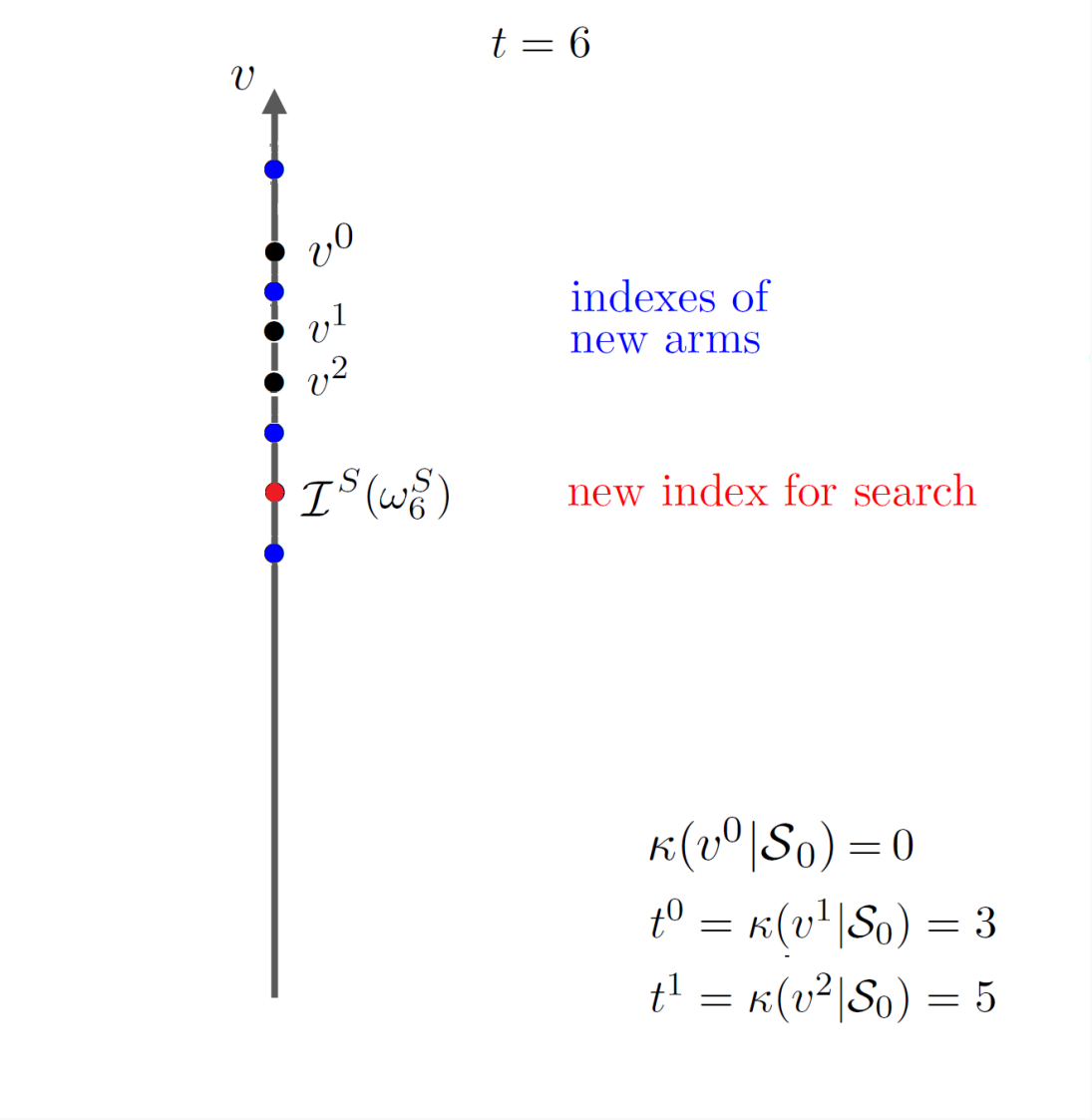



$$\kappa(v^0|\mathcal{S}_0) = 0$$

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

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

# Proof of Lemma 1



# 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 $\kappa(v^{i+1})-1$:

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

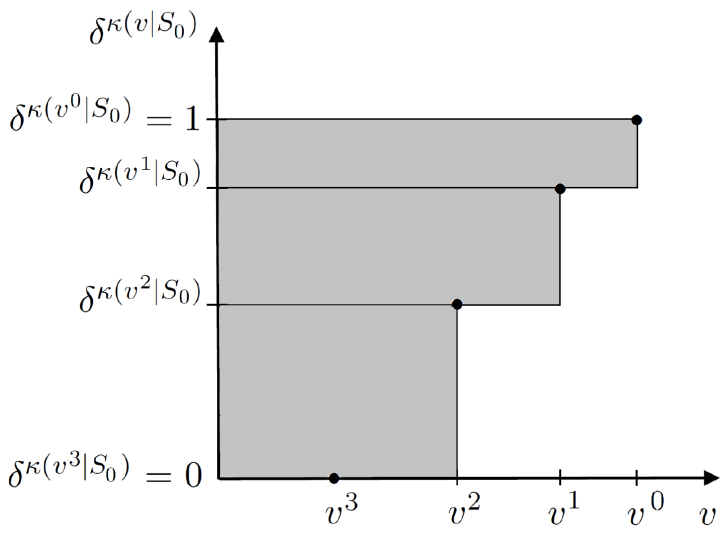
- Same true if multiple alternatives and/or search have index equal to $v^i$ at $\kappa(v^i)$

# 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: Step 2

$\mathcal{V}(\mathcal{S}_0)$ solves Bellman 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)
- decomposition of overall problem into collection of binary problems where choice is btw single alternative (possibly search) and auxiliary alternatives with fixed payoff.

# 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

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

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

Similarly, for physical alternative in state $\omega^P$:

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

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

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

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

  - 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