Robustness in Search and Experimentation

Suraj Malladi

Northwestern, Kellogg

Robustness Summer School

August 26th, 2026

Decision-making under ignorance

Usual motivation for models of DMs with robust objectives:

- environment is complex
- little is known
- do well regardless of what happens

Static approach: maximizes guaranteed payoff for a one-off decision

E.g., designing an institution or policy that cannot soon be revised

Ignorance is temporary

Agents can often learn from past choices and adapt

E.g., setting an initial price for a new product

Need not pick best static action if choice affects learning

Dynamic problem: maximize guaranteed *continuation payoff*

Starting point: single-agent search

1. Pick an item from a consideration set
2. Learn its payoff
3. Update about payoffs to unsearched items

E.g., search for products, ideas, prices, partners…

Step 1 is sometimes assumed away (*undirected search*)
Step 3 is often assumed away (*independent payoffs*)
Step 1+3 rarely assumed jointly

But *directed search and learning* is ubiquitous

Directed search and learning

Weitzman (**1979**): "*It appears plausible that other things being equal it would be better to open a box whose reward is highly correlated with other rewards because this adds a positive informational externality. But translating such an effect into a simple search rule seems difficult except in the most elementary cases.*"

Francetich and Kreps (**2020**): "*except in very special cases, such [problems are] unsolvable—either analytically or numerically*"

Common approaches: behavioral, approximations, limited foresight…

Robust approach is tractable! (and rational!)

Agenda

1. A framework for robust search and experimentation

2. Applications
    1. Motivate robust approach
    2. Solve the model
    3. Suggest open questions

3. Dynamic consistency

A General Framework

Malladi (2023)

Preliminaries

One agent

A search space $S \subset \mathbb{R}_+$

Persistent state, $q: S \to \mathbb{R}_+$

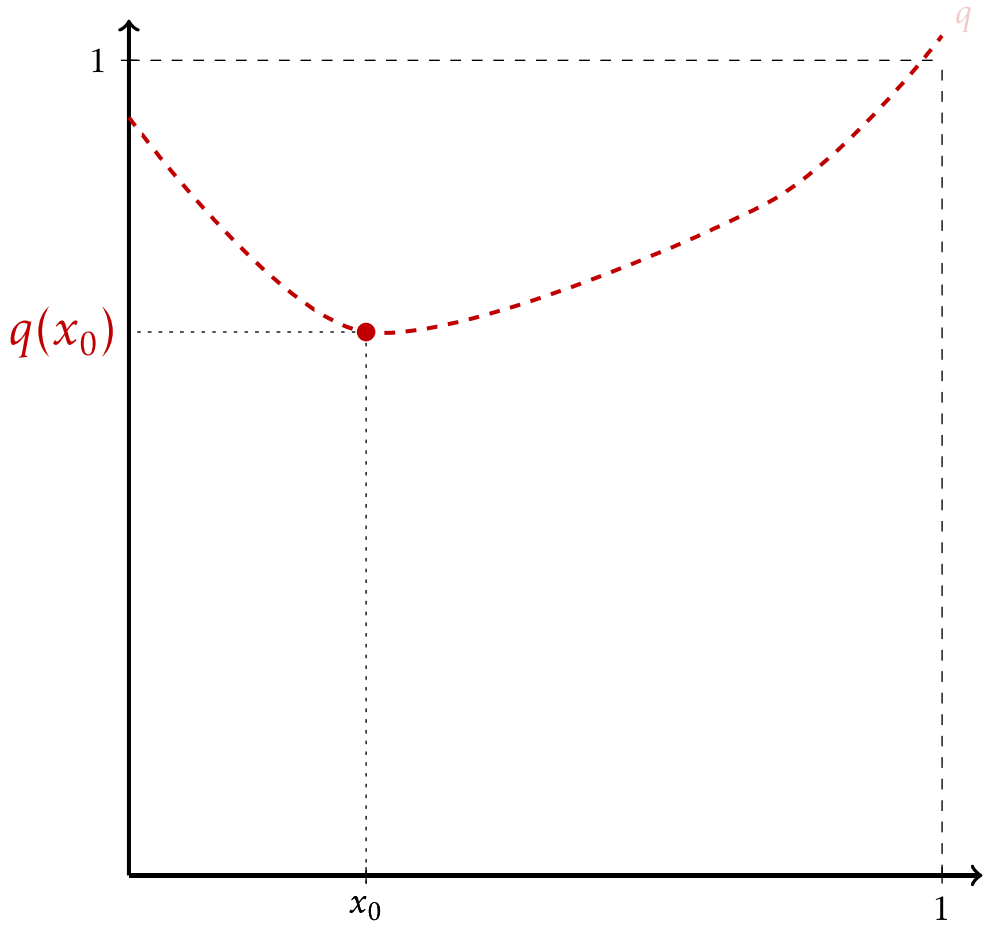
Initially, q is unknown: only know that $q \in \Omega$

Actions

x_t: agent's time t action

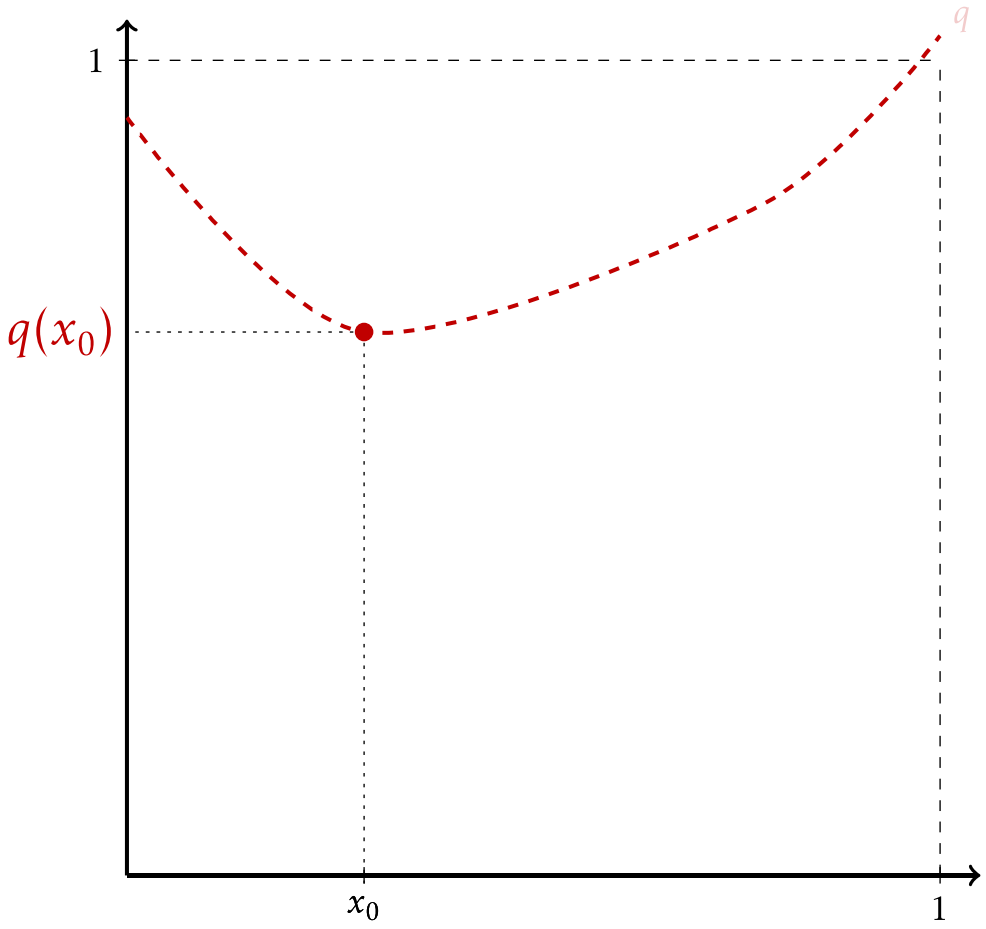
$x_t \in S$: learns $q(x_t)$

$x_t = \emptyset$: search ends



Learning

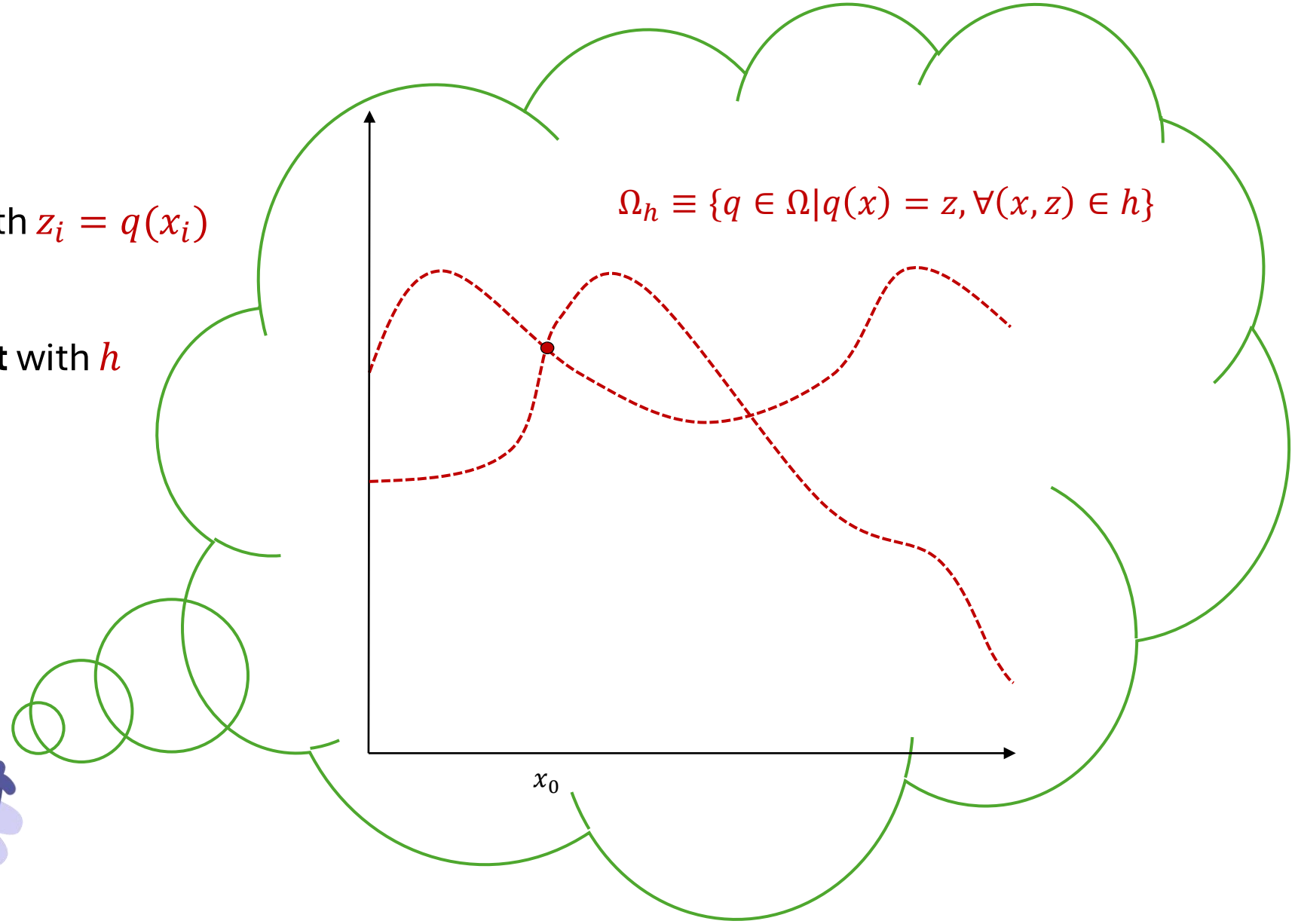
$h_t = (x_i, z_i)_{i=0}^{t-1}$: history with $z_i = q(x_i)$



Learning

$h_t = (x_i, z_i)_{i=0}^{t-1}$: history with $z_i = q(x_i)$

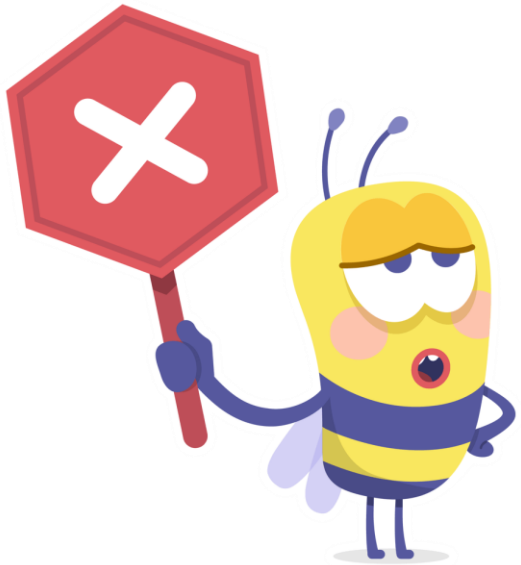
$\Omega_h \subset \Omega$: states **consistent** with h

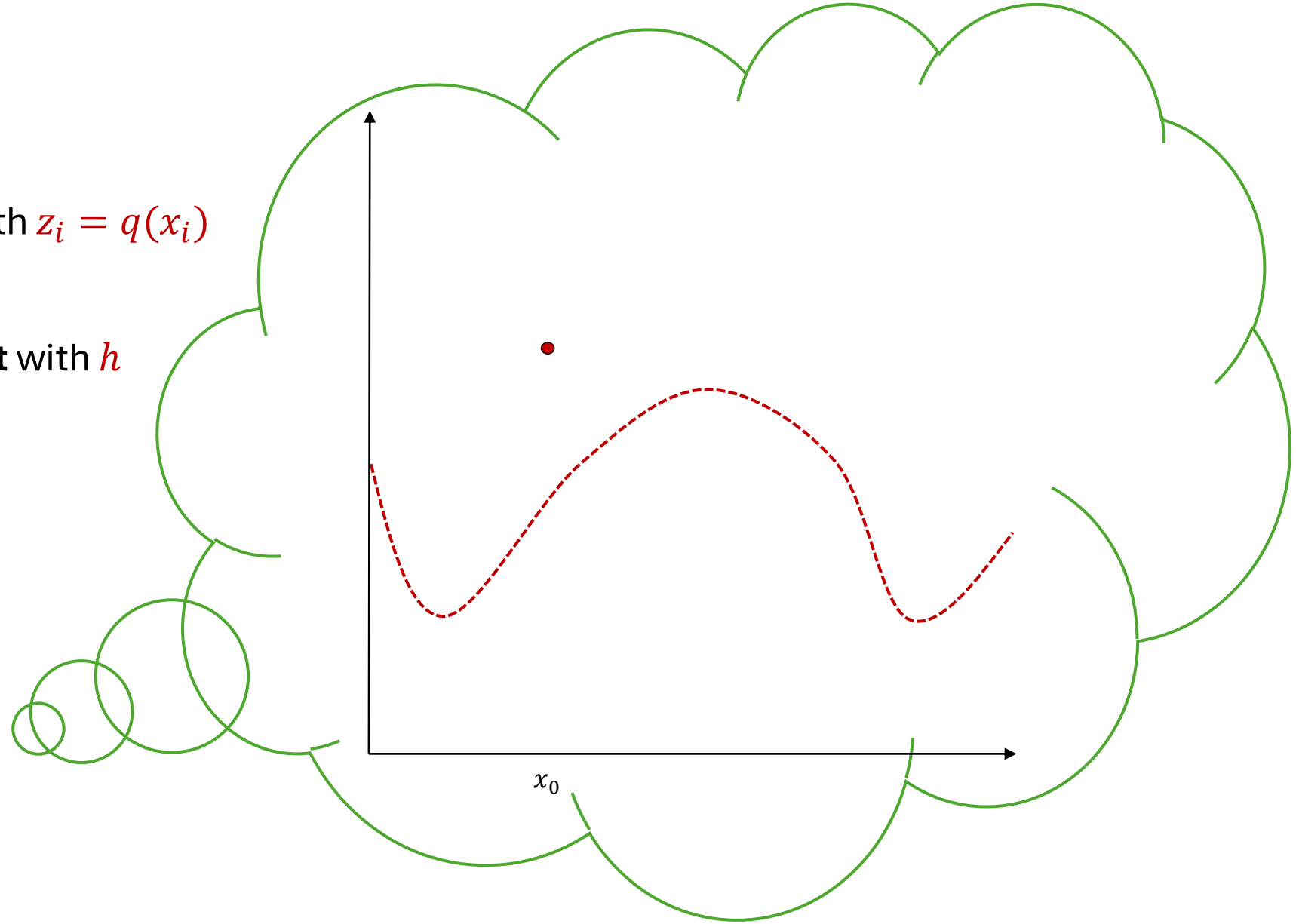


Learning

$h_t = (x_i, z_i)_{i=0}^{t-1}$: history with $z_i = q(x_i)$

$\Omega_h \subset \Omega$: states **consistent** with h





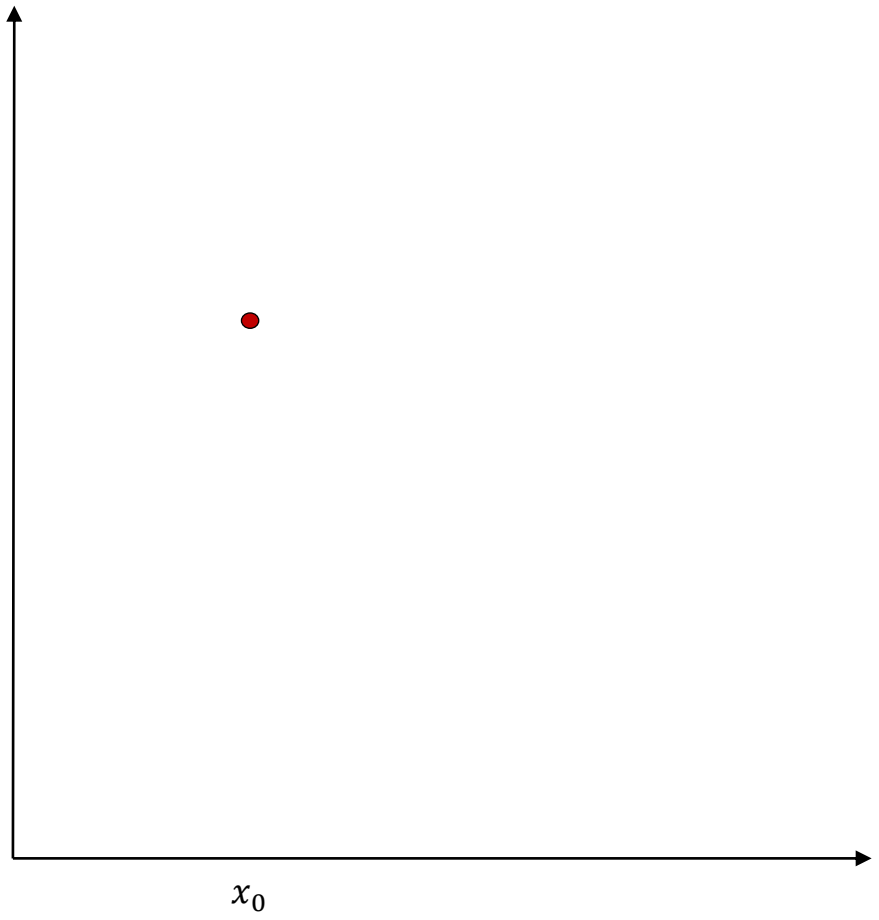
Learning

$h_t = (x_i, z_i)_{i=0}^{t-1}$: history with $z_i = q(x_i)$

$\Omega_h \subset \Omega$: states **consistent** with h

H: histories h where Ω_h nonempty

$\boldsymbol{h} \in H$: **path** is an infinite history



Strategies

A **strategy** $\sigma \in \Sigma$ is any map $\bigcup_{t\in\mathbb{N}} H_t \to S \cup \{\emptyset\}$

History h_t, strategy $\sigma \in \Sigma$, and state $q \in \Omega_{h_t}$ induce **path** $\boldsymbol{h}(\sigma, q, h_t)$

Payoffs

$\Pi(\sigma, q, h_t)$ denotes the **continuation payoff** to σ at h_t when state is q

At h_t, strategy σ has **guaranteed continuation payoff**:

$$\min_{q \in \Omega_{h_t}} \Pi(\sigma, q, h_t)$$

Ex ante optimality

σ^* is **ex ante optimal** if it maximizes guaranteed continuation payoff at $h = h_0$:

$$\sigma^* \in \underset{\sigma \in \Sigma}{\operatorname{argmax}} \min_{q \in \Omega} \Pi(\sigma, q, h_0)$$

Sequential optimality

For a strategy $\sigma \in \Sigma$, history h_t is **reachable** if…

1. it is h_0,
2. or h_t is the t-truncation of $\boldsymbol{h}(\sigma, q, h_0)$ for some $q \in \Omega$.

σ^* is **sequentially optimal** if it maximizes guaranteed payoff at all reachable h_t:

$$\sigma^* \in \underset{\sigma \in \Sigma}{\operatorname{argmax}} \min_{q \in \Omega_{h_t}} \Pi(\sigma, q, h_t)$$

Applications

Applications vary by assumptions on state space Ω and payoffs Π

Price experimentation:

- Ω is space of demand curves
- Π is sum of discounted flow profits

Search and learning:

- Ω is some space of continuous curves
- Π is best discovery minus sum of search costs

Price Experimentation

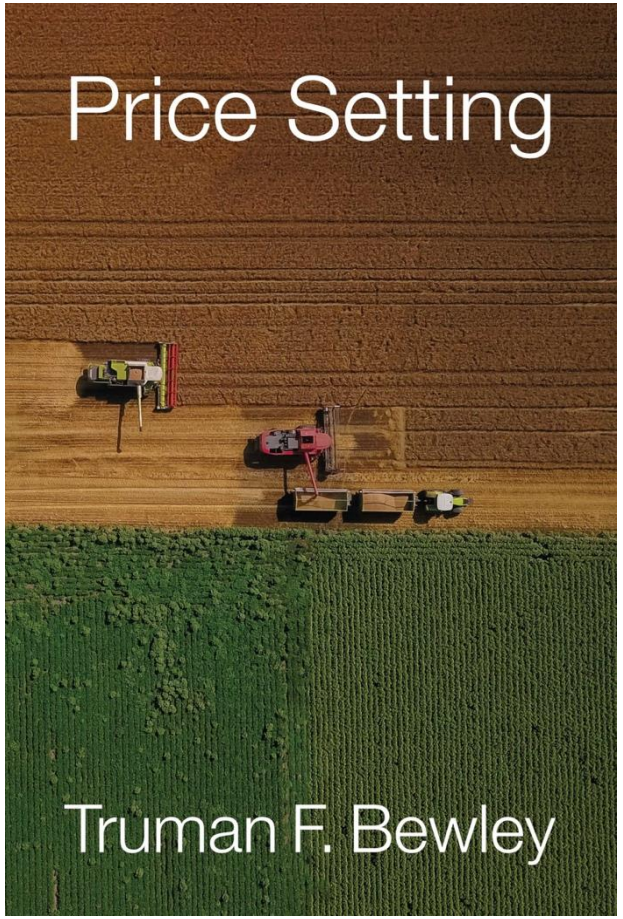
Malladi (2025)

Question

How do firms set prices when they don't know their demand curve?

- Static framing: what price should we set once and for all?
  - Bergemann and Schlag (2008, 2011)
- But firms change prices over time to learn demand
  - "fit demand curve to data, move toward optimal price"
- How do they do this?

What firms say they do



Monopolistically competitive firms:

slowly **raise prices** over time, seldom reduce them

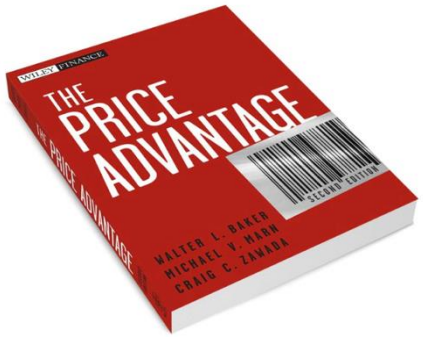
act as if **demand is kinked**:
"*demand responds little to price reduction…[but] a price increase might…reduce demand sharply*."

Consultants/VCs say similar things

"[Many executives] assert that pricing is a zero-sum game: if they raise prices, the lost volume will negate the effect of higher margins; if they lower prices, the additional volume will not be enough to compensate…"

BCG Perspectives (2003)

Suggest learning demand by raising prices over time:



10 · 5 · 20 RULE

Value = 10X Price
Raise price 5%
Until 20% Push Back

Why might a firm behave this way?

Model of a cautious firm learning its demand

- Long-lived firm sets a price each period and sees demand
- Maximizes guaranteed sum of discounted flow profits

Multiarm bandit with correlated arms

Each price is an arm, and reward is that period's profit

Demand at one price informative about demand at other prices

Models where firms maximize expected profits are intractable

Optimal experimentation path for firm with long planning horizon?

"*Two approaches have appeared to the question of experimentation in the face of a random demand curve…One involves formulating an infinite horizon model in which attention turns to the limiting expectation…A second approach is restricted to two-period models*"

Mirman, Samuelson, Urbano (1993)

How do firms misprice when they don't learn the monopoly price?

"*When adequate learning does not obtain one cannot understand the long-run outcome independently of the priors or the adjustment process by which it was reached. An entirely new kind of analysis is called for in these cases: in order to determine the nature of the long-run behaviour of the agent, one needs to characterise the optimal learning strategy. Unfortunately this is possible analytically only in very simple learning problems*"

Aghion, Bolton, Harris, Jullien (1991)

Outline

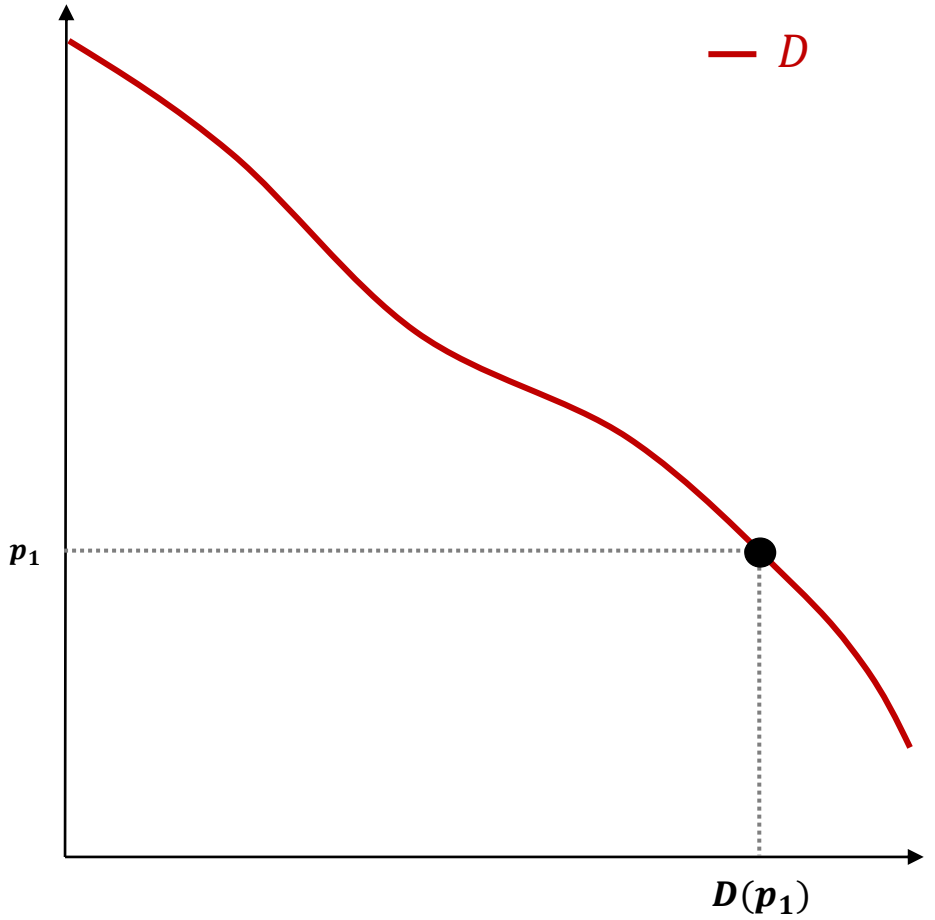
1. Model
2. Special case
3. General case
4. Discussion

Actions and learning

At $t = 0, 1, 2 \ldots$ the agent:

sets price p_{t+1}

observes demand $D(p_{t+1})$



Payoffs

Let $\delta \in (0,1)$; **continuation profit** to σ at h_t is

$$\Pi(\sigma, D, h_t) \equiv \sum_{i=t}^{\infty} \delta^i \cdot p_{i+1}^{\boldsymbol{h}(\sigma,D,h_t)} \cdot D(p_{i+1}^{\boldsymbol{h}(\sigma,D,h_t)})$$

At h_t, strategy σ has **guaranteed continuation profit**:

$$\min_{d \in \Omega_{h_t}} \Pi(\sigma, d, h_t)$$

Assumptions on Ω

1. Each $d \in \Omega$ is strictly decreasing at any p where $d(p) > 0$

2. Ω is uniformly equicontinuous

3. Inverse demand Ω^{-1} is uniformly equicontinuous

4. Ω is closed under the supremum norm

Outline

1. Model
2. Special case
3. General case
4. Discussion

Special case: firm knows little about demand

Let $p_0, q_0 > 0$ and $\underline{\tau}, \overline{\tau} < 0$

Assumption 1: Ω is set of demand curves $d: \mathbb{R}_+ \to \mathbb{R}_+$ where

1. $d(p_0) \geq q_0$,
2. $\underline{\tau} \leq \frac{d(x)-d(y)}{x-y} \leq \overline{\tau}$ for all $x \neq y$ where $d(x), d(y) > 0$.

Kinked pricing strategy

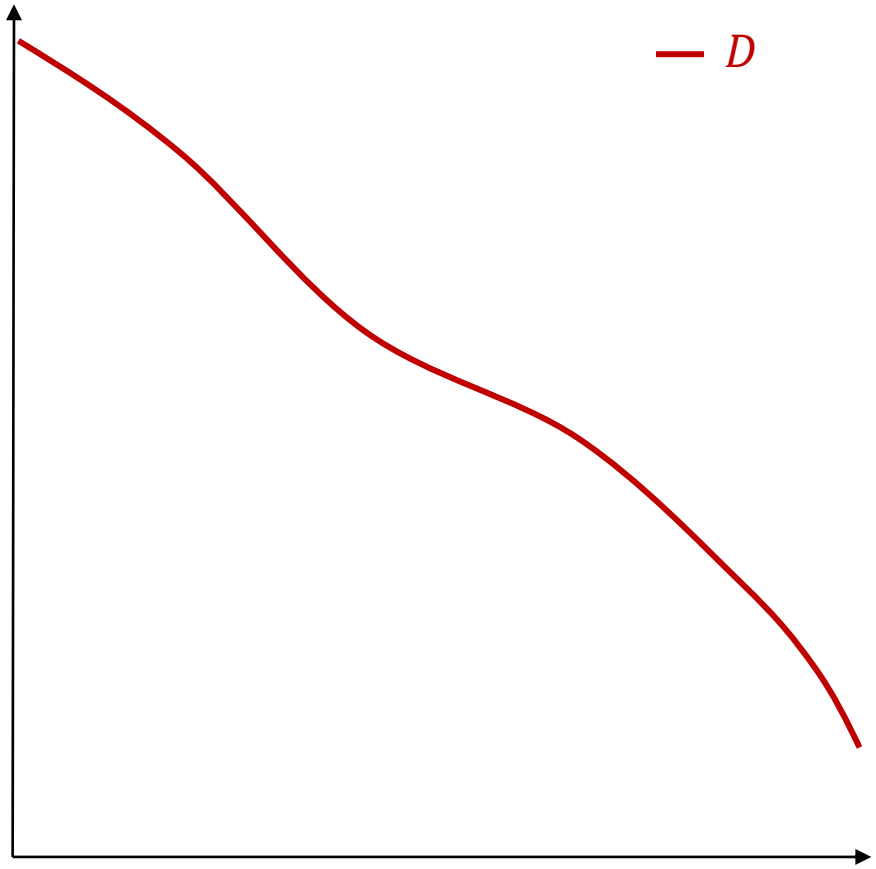
At any history $h_t = (p_i, q_i)_{i=1}^{t}$, let

$$d_{h_t}(p) \equiv q_t + \min\{\underline{\tau}(p - p_t), \overline{\tau}(p - p_t)\}$$

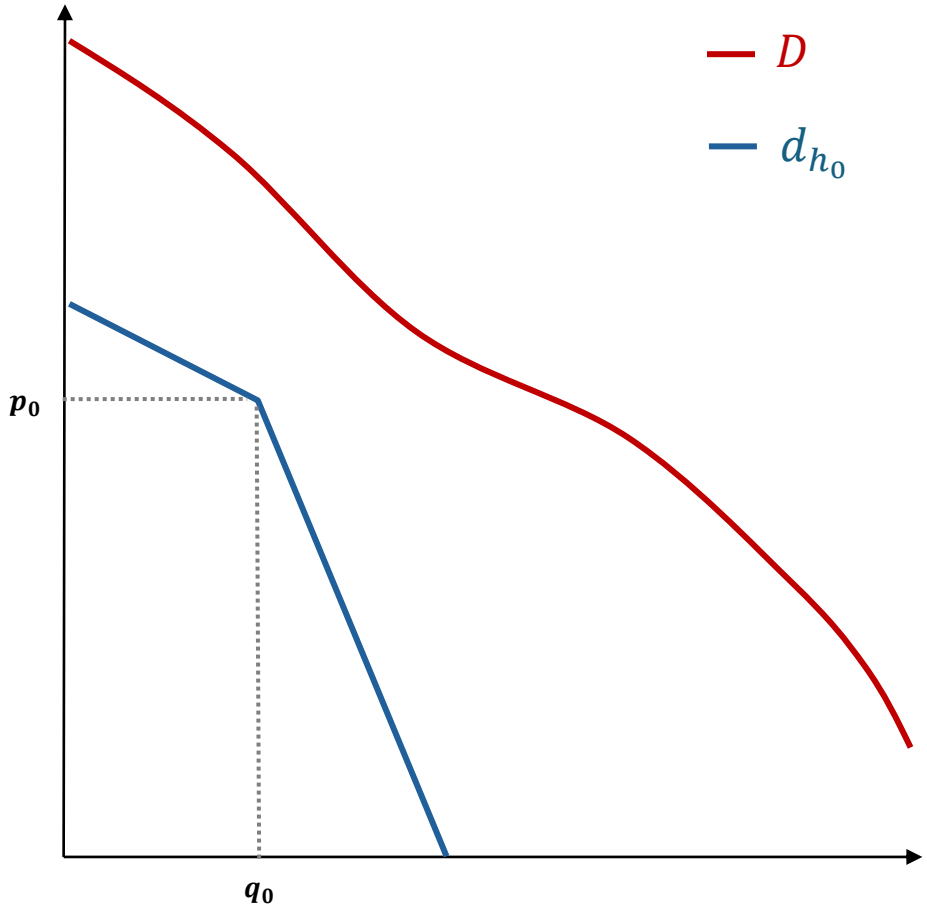
Let the **kinked pricing strategy** σ_k be defined as

$$\sigma_k(h_t) \equiv \underset{p}{\operatorname{argmax}}\, p \cdot d_{h_t}(p)$$

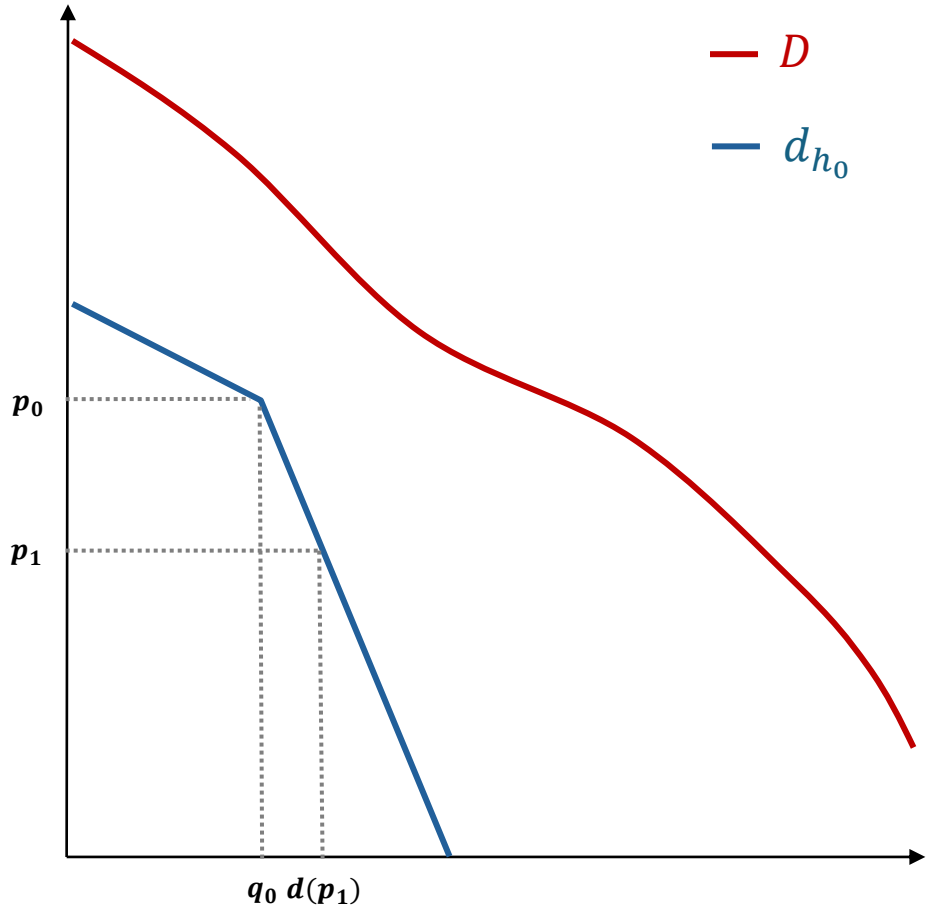
Kinked pricing strategy



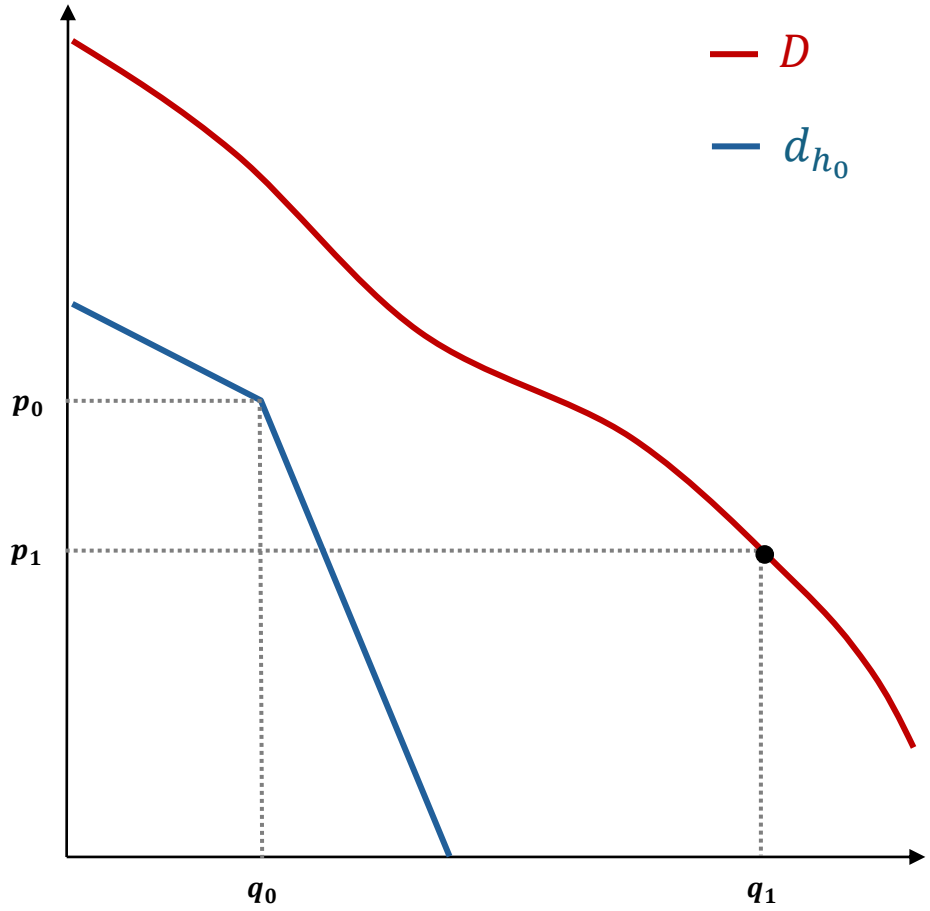
Kinked pricing strategy



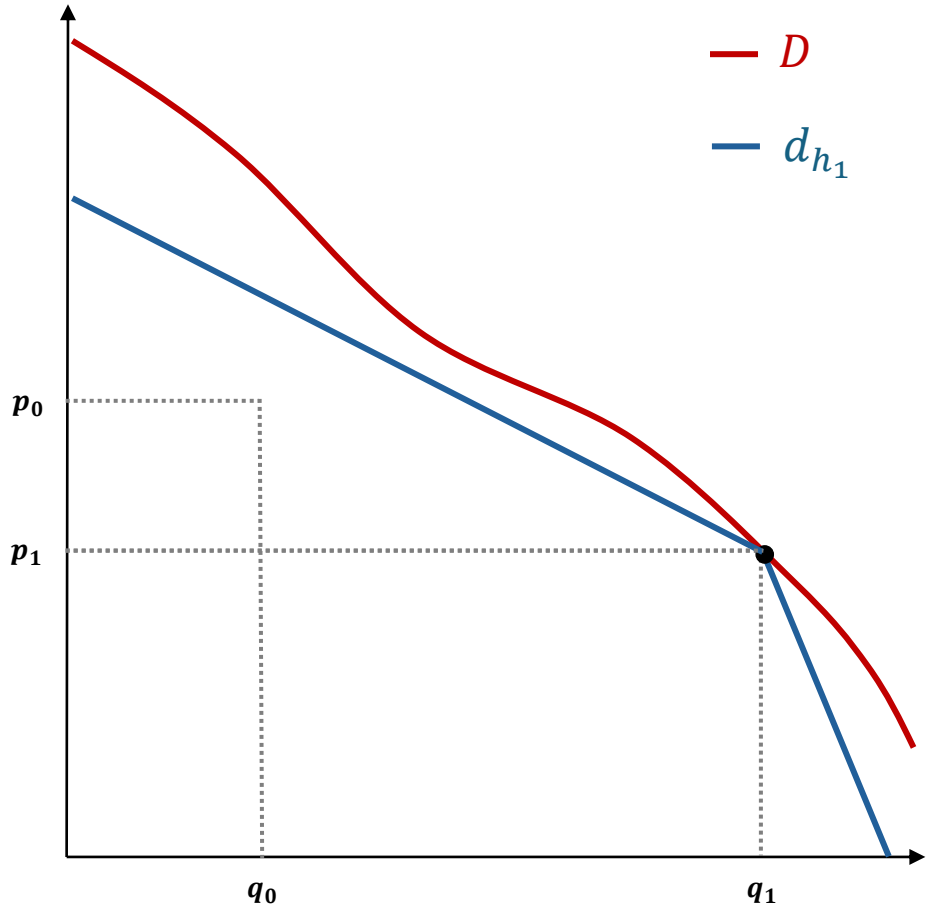
Kinked pricing strategy



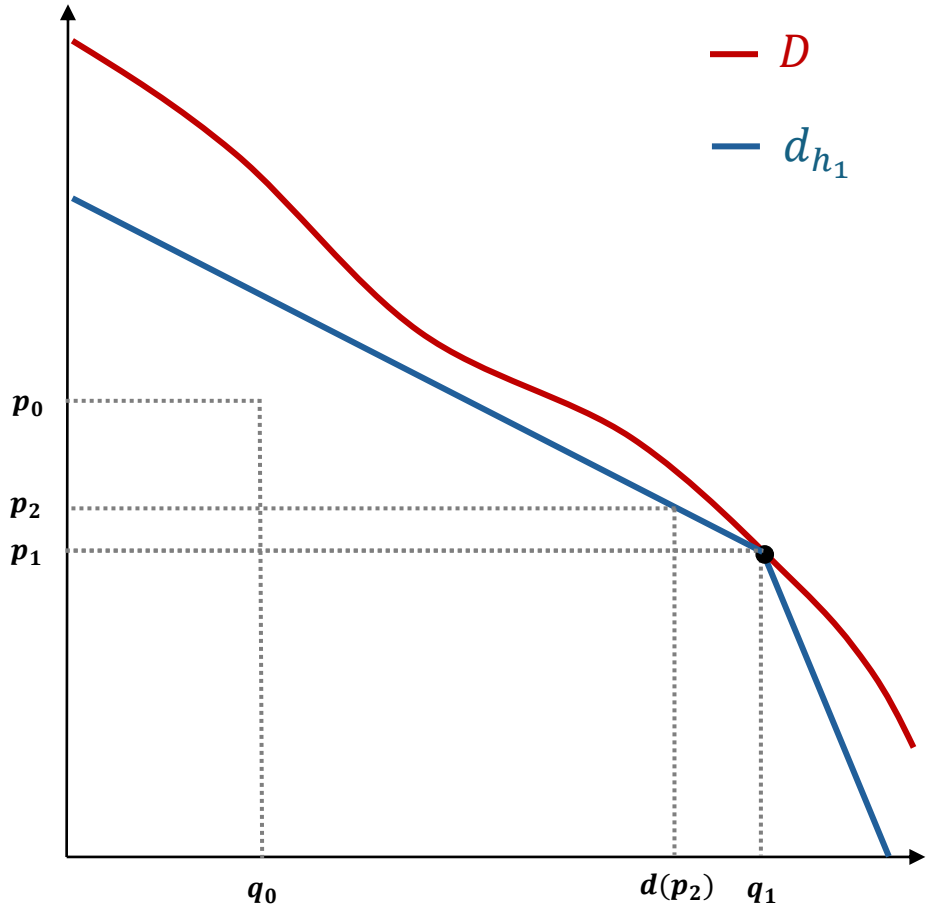
Kinked pricing strategy



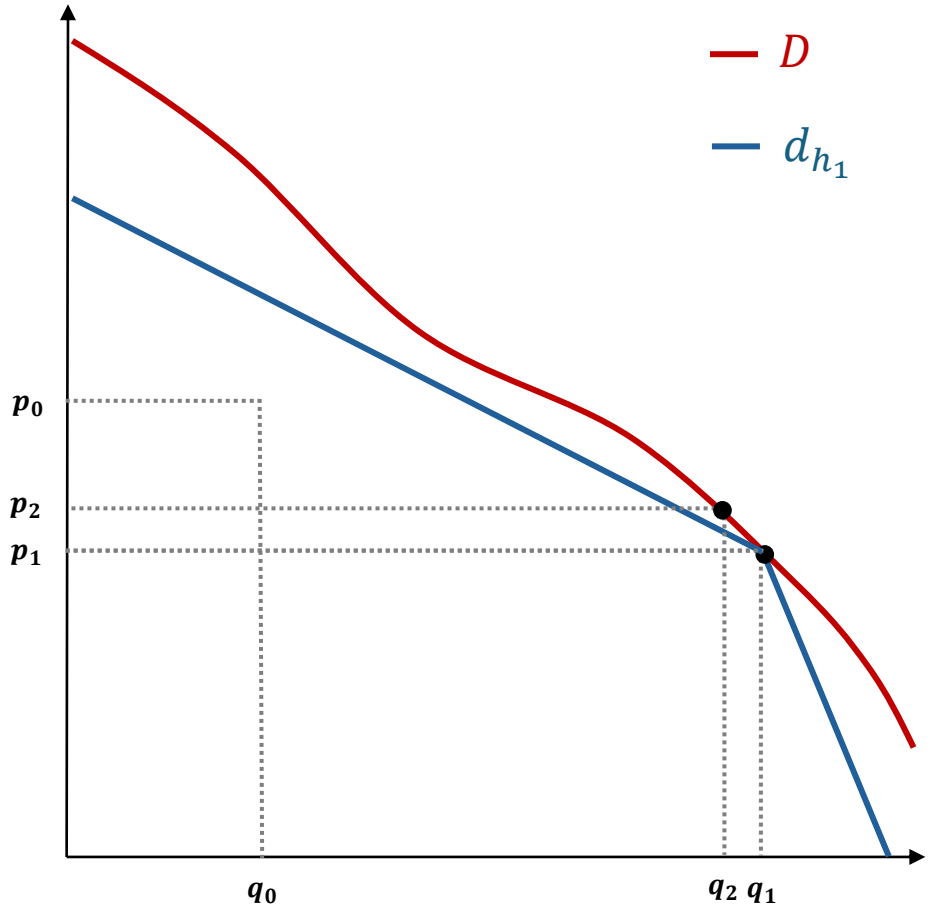
Kinked pricing strategy



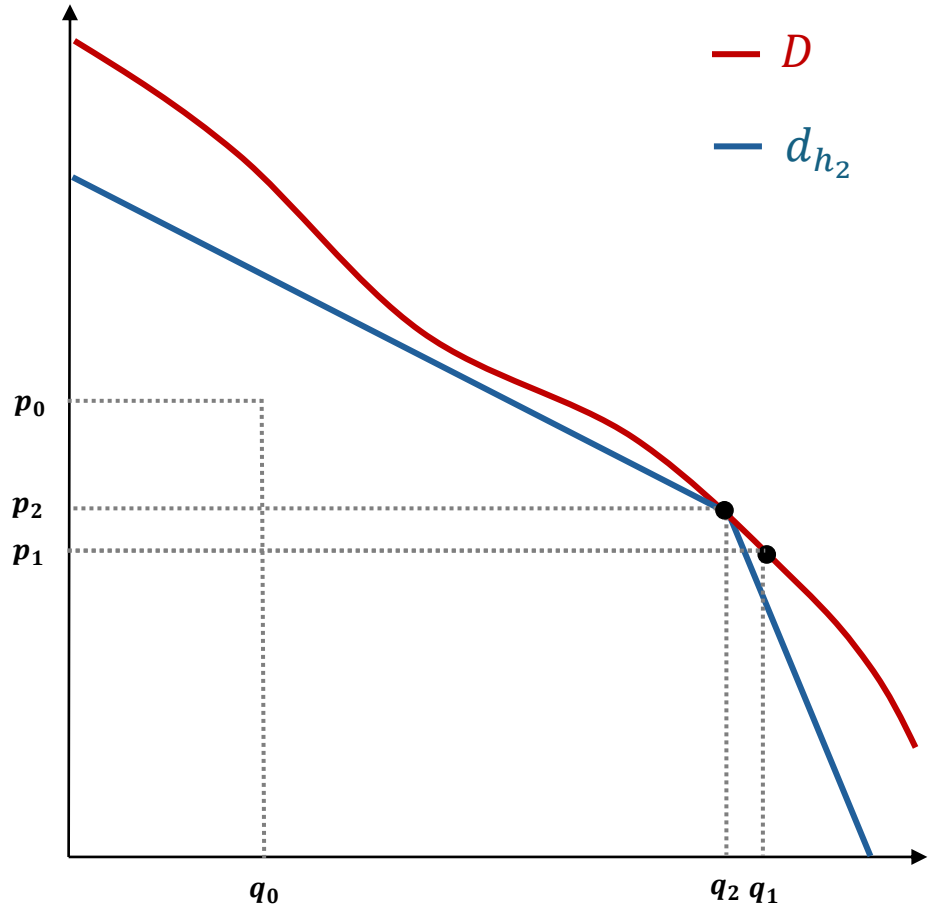
Kinked pricing strategy



Kinked pricing strategy



Kinked pricing strategy (memoryless)



Pricing path

Proposition 1: The kinked pricing strategy σ_k is seq. optimal.

Uniqueness

The optimal strategy is **essentially unique** if for any two optimal strategies σ_1, σ_2 and any $d \in \Omega$, both strategies lead down the same path, i.e.,

$$\boldsymbol{h}(\sigma_1, d, h_0) = \boldsymbol{h}(\sigma_2, d, h_0)$$

Pricing path

Proposition 1: The kinked pricing strategy σ_k is seq. optimal.

Proposition 2: The seq. optimal strategy is essentially unique.

Rising prices and profits

Prices always rise under strategy σ if, for any $d \in \Omega$ and along the path $\boldsymbol{h} = \boldsymbol{h}(\sigma, d, h_0)$, $p_i^{\boldsymbol{h}}$ is increasing in i

Profits always rise under strategy σ if, for any $d \in \Omega$ and along the path $\boldsymbol{h} = \boldsymbol{h}(\sigma, d, h_0)$, flow profit $p_i^{\boldsymbol{h}} \cdot d(p_i^{\boldsymbol{h}})$ is increasing in i

Pricing path

Proposition 1: The kinked pricing strategy σ_k is seq. optimal.

Proposition 2: The seq. optimal strategy is essentially unique.

Proposition 3: Prices and profits always rise under σ_k.

Bewley (2025)

When price setters report kinked demand curves, should we:

"*provisionally accept as an empirical regularity price setters' beliefs…To go beyond this disappointing stance…[one] needs to verify that price setters' beliefs are accurate*"

Prop 2 + 3: as if demand is kinked + downward price rigidity

Long run

For any $d \in \Omega$ and optimal σ, prices along $\boldsymbol{h}(\sigma, d, h_0)$ converge to a **long-run price**, p_d^∞

Denote the smallest **informed monopoly price** by

$$p_d^* \equiv \underset{p}{\operatorname{argmax}}\, p \cdot d(p)$$

There is **underpricing** if $p_d^\infty < p_d^*$

There is **price rigidity** if $\sigma_k(h_1) = \sigma_k(h_0)$

Long run

Proposition 1: The kinked pricing strategy σ_k is seq. optimal.

Proposition 2: The seq. optimal strategy is essentially unique.

Proposition 3: Prices and profits always rise under σ_k.

Proposition 4: Generically, there is price rigidity or underpricing under σ_k.

Long run

Proposition 1: The kinked pricing strategy σ_k is seq. optimal.

Proposition 2: The seq. optimal strategy is essentially unique.

Proposition 3: Prices and profits always rise under σ_k.

Proposition 4: Generically, there is price rigidity or underpricing under σ_k.

Proposition 5: $\frac{p_d^\infty \cdot d(p_d^\infty)}{p_d^* \cdot d(p_d^*)} \geq \frac{4 \cdot \underline{\tau} \cdot \overline{\tau}}{(\underline{\tau}+\overline{\tau})^2}$.

Why profit guarantee maximization?

Firm may want to know what it can guarantee itself

- Proposition 5 answers this
- If guarantee is high, σ_k is attractive
- If guarantee is low, value in learning more about elasticities

Proof ideas: passive learning

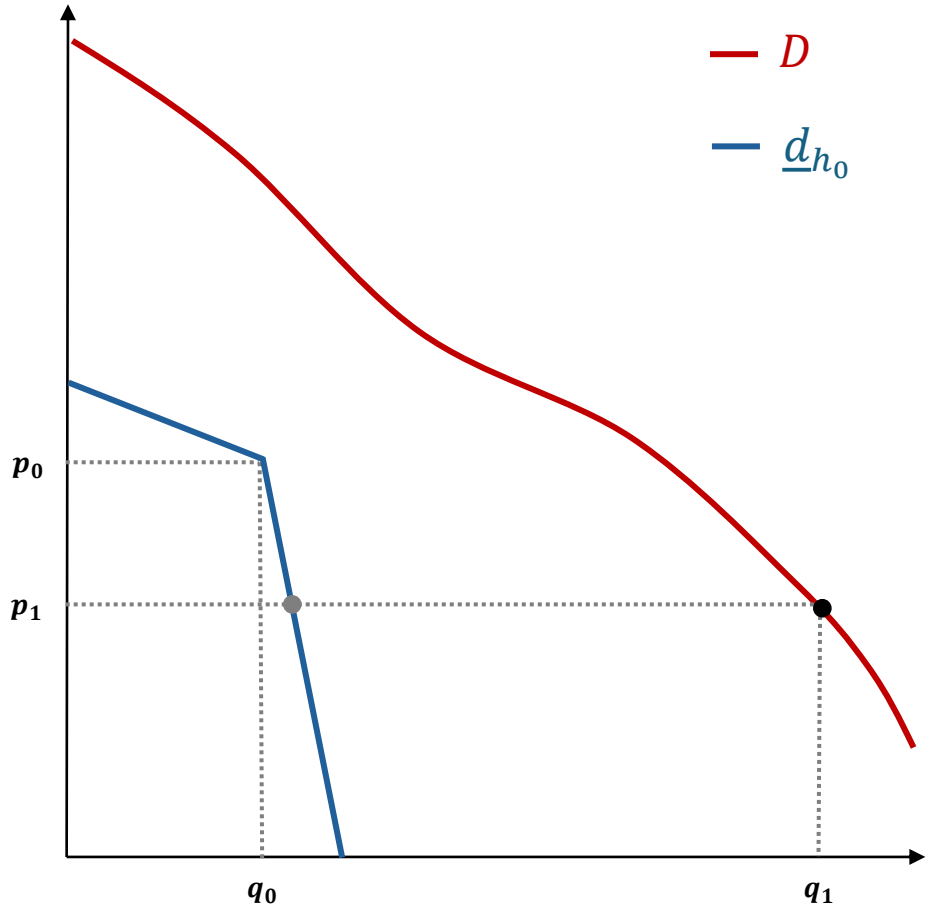
At any history h_t, denote the lower envelope of Ω_{h_t} as

$$\underline{d}_{h_t}(p) \equiv \inf_{d \in \Omega_{h_t}} d(p)$$

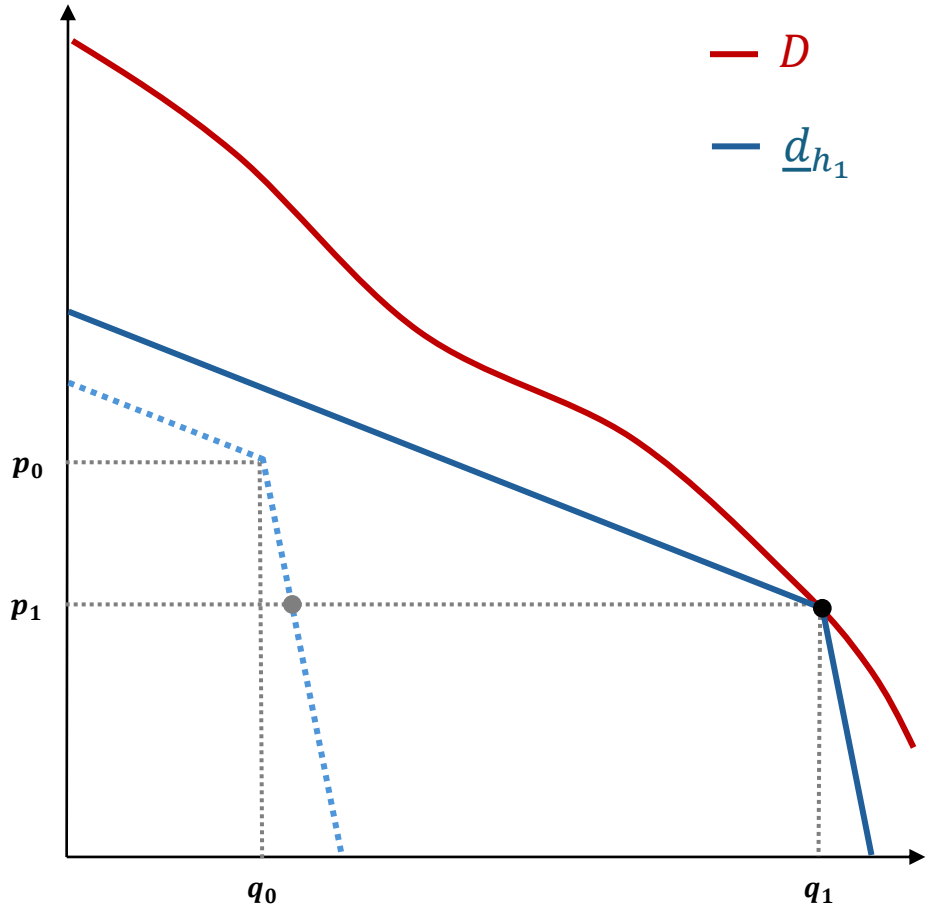
Define the **myopic strategy** $\underline{\sigma}$ as

$$\underline{\sigma}(h_t) \equiv \min \operatorname*{argmax}_p p \cdot \underline{d}_{h_t}(p)$$

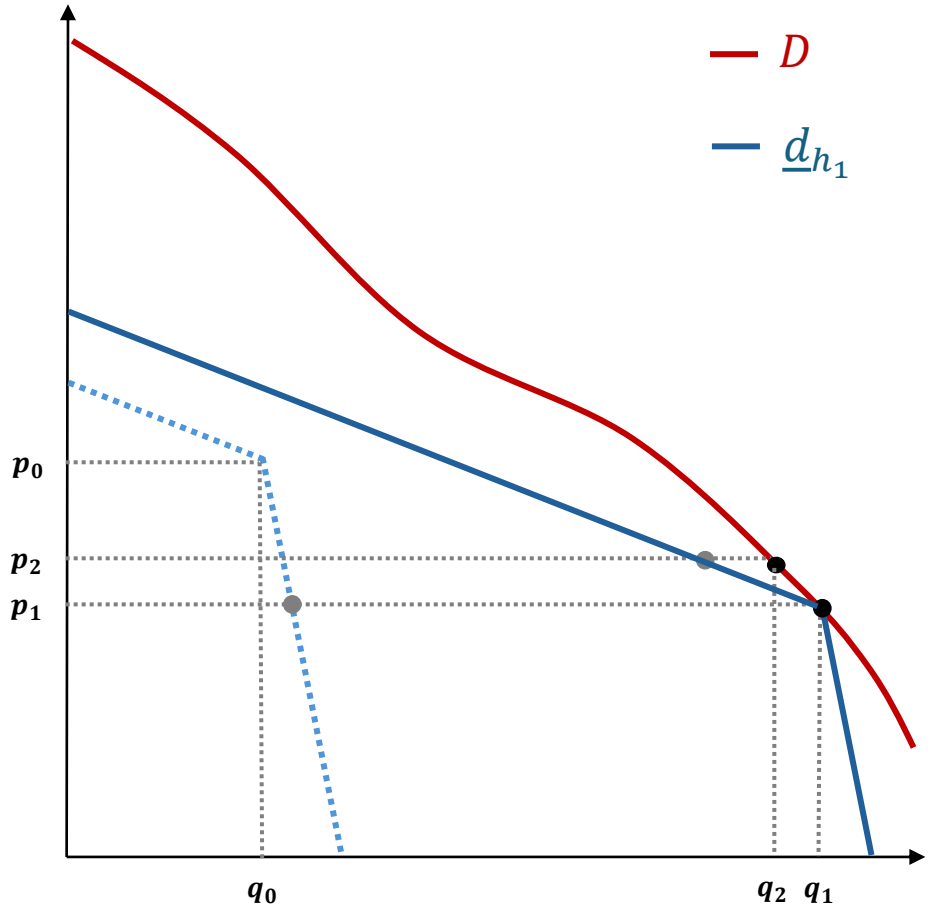
Lemma 1: Prices/profits always rise under $\underline{\sigma}$



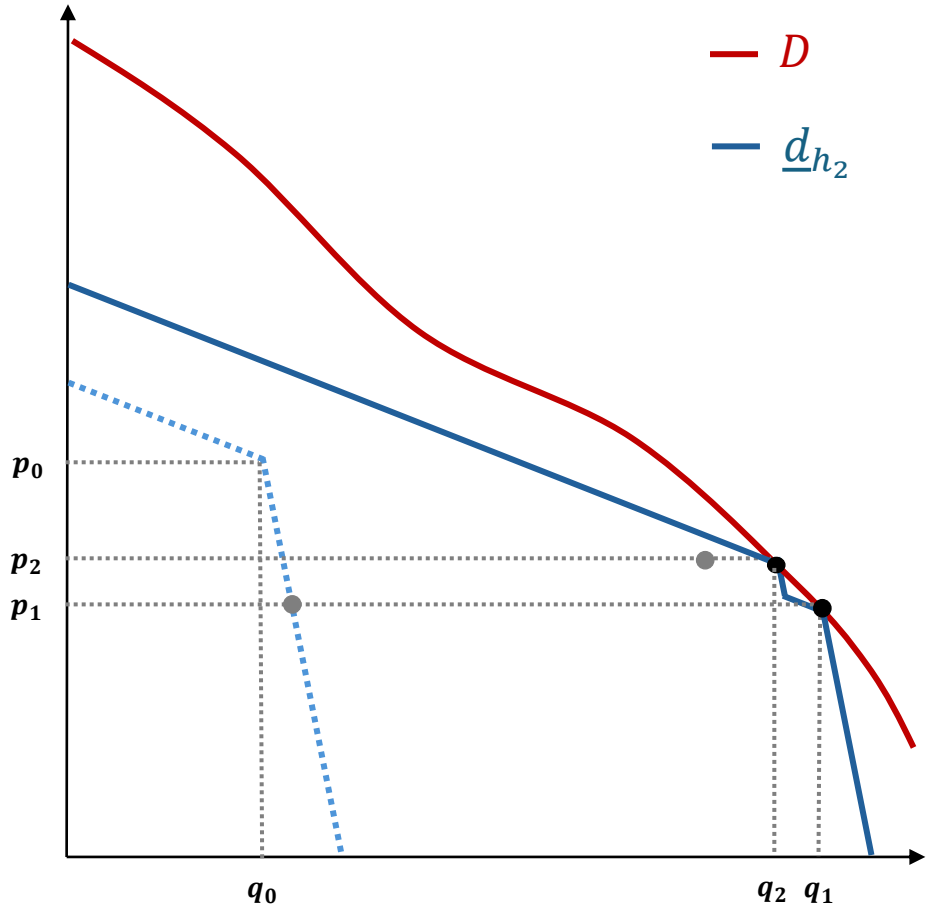
Lemma 1: Prices/profits always rise under $\underline{\sigma}$



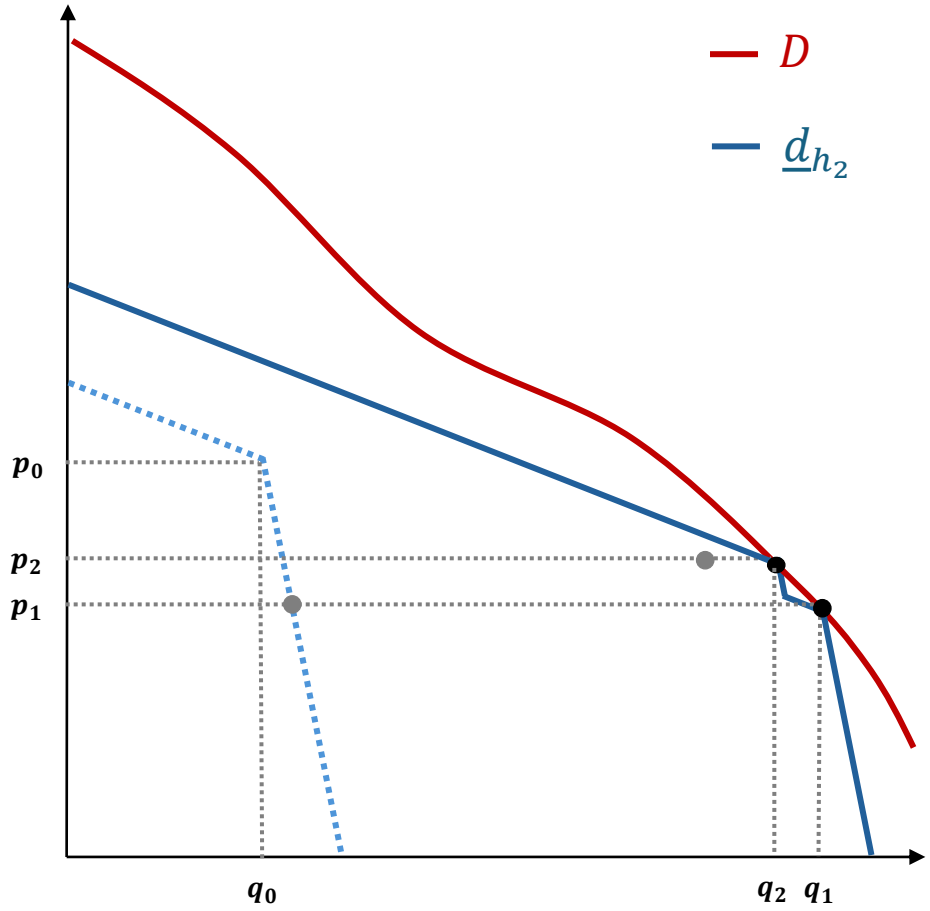
Lemma 1: Prices/profits always rise under $\underline{\sigma}$



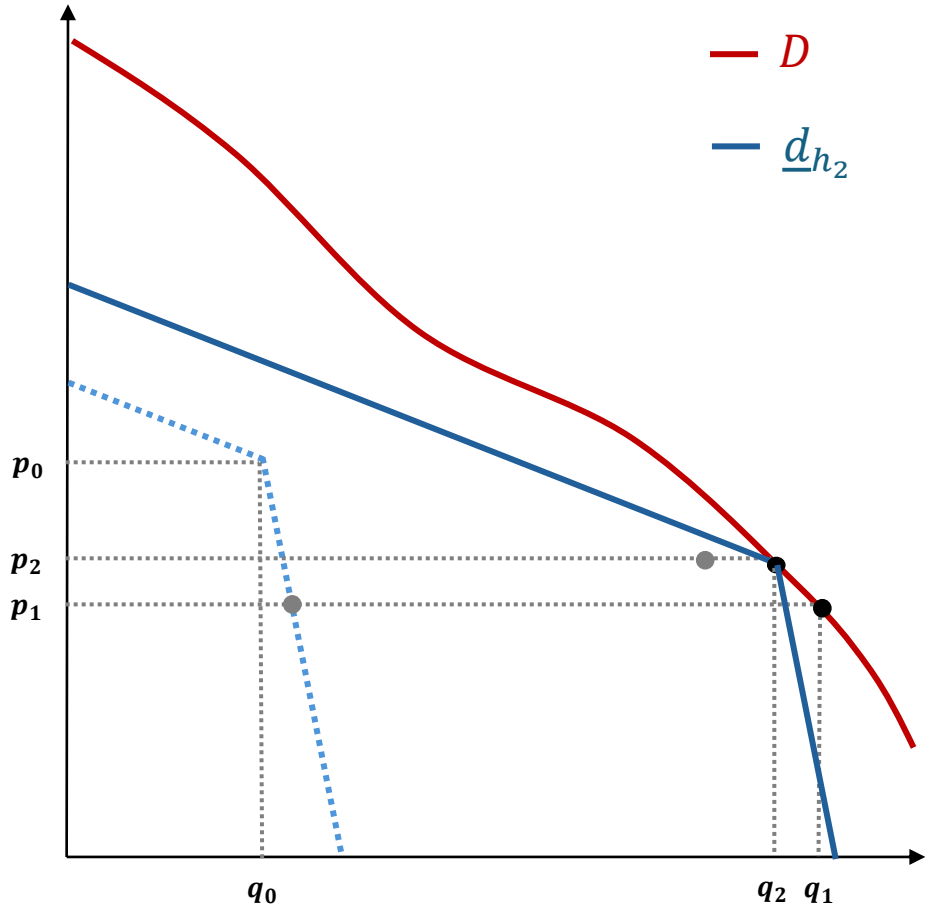
Lemma 1: Prices/profits always rise under $\underline{\sigma}$



Lemma 2: $\underline{\sigma} = \sigma_k$ on path



Lemma 2: $\underline{\sigma} = \sigma_k$ on path



Lemma 3: $\underline{\sigma}$ is sequentially optimal

For any history h_t and strategy σ, $\Pi(\sigma, \underline{d}_{h_t}, h_t) \leq \Pi(\underline{\sigma}, \underline{d}_{h_t}, h_t)$

For any history h_t and demand $d \in \Omega_{h_t}$, $\Pi(\underline{\sigma}, \underline{d}_{h_t}, h_t) \leq \Pi(\underline{\sigma}, d, h_t)$
by Lemma 1

Therefore, $\underline{\sigma}$ is sequentially optimal

A rationalization of passive learning

"*pricing policies used in various industries are often myopic, putting no particular emphasis on active learning…This conventional approach is not truly optimal, because it ignores the fact that the estimate-and-optimize cycle will be repeated in the future*"

Harrison, Keskin, Zeevi (2012)

Outline

1. Model
2. Special case
3. General case
4. Discussion

When are myopic strategies optimal?

Ω is ***downward closed*** if $\underline{d}_{h_t} \in \Omega_{h_t}$ for all reachable h_t under $\underline{\sigma}$

Ω is ***effectively d.c.*** if for all reachable h_t under $\underline{\sigma}$, for some $d \in \Omega_{h_t}$

$$\max_p p \cdot \underline{d}_{h_t}(p) = \max_p p \cdot d(p)$$

When are myopic strategies optimal?

Theorem 1

If state space Ω is effectively downward closed, then some myopic strategy is sequentially optimal.

Conversely, if Ω is not effectively downward closed, then for a sufficiently patient firm, the myopic strategy is not sequentially optimal.

Proof sketch for Theorem 1

If Ω is effectively downward closed, same saddle point argument as before.

If Ω not effectively downward closed, consider history h where condition fails.

Let $m = \inf_{d \in \Omega_h} \max_p p \cdot d(p)$

By Arzela-Ascoli theorem $m > \max_p p \cdot \underline{d}_h(p) = \underline{m}$

Experimenting on fine enough grid reveals a price where flow profit $> \underline{m}$

Applications: concave demand

Assumption 2: Ω is the set of **concave** demand curves $d: \mathbb{R}_+ \to \mathbb{R}_+$ that satisfy Assumption 1

Lower envelope of Ω_h is also concave and satisfies Assumption 1

Downward closed → myopic strategy is seq. optimal

Locally informative (see paper) → prices always rise

Applications: convex demand

Assumption 3: Ω is the set of **convex** demand curves $d: \mathbb{R}_+ \to \mathbb{R}_+$ that satisfy Assumption 1

Lower envelope of Ω_h is also not convex

Effectively downward closed → myopic strategy is seq. optimal

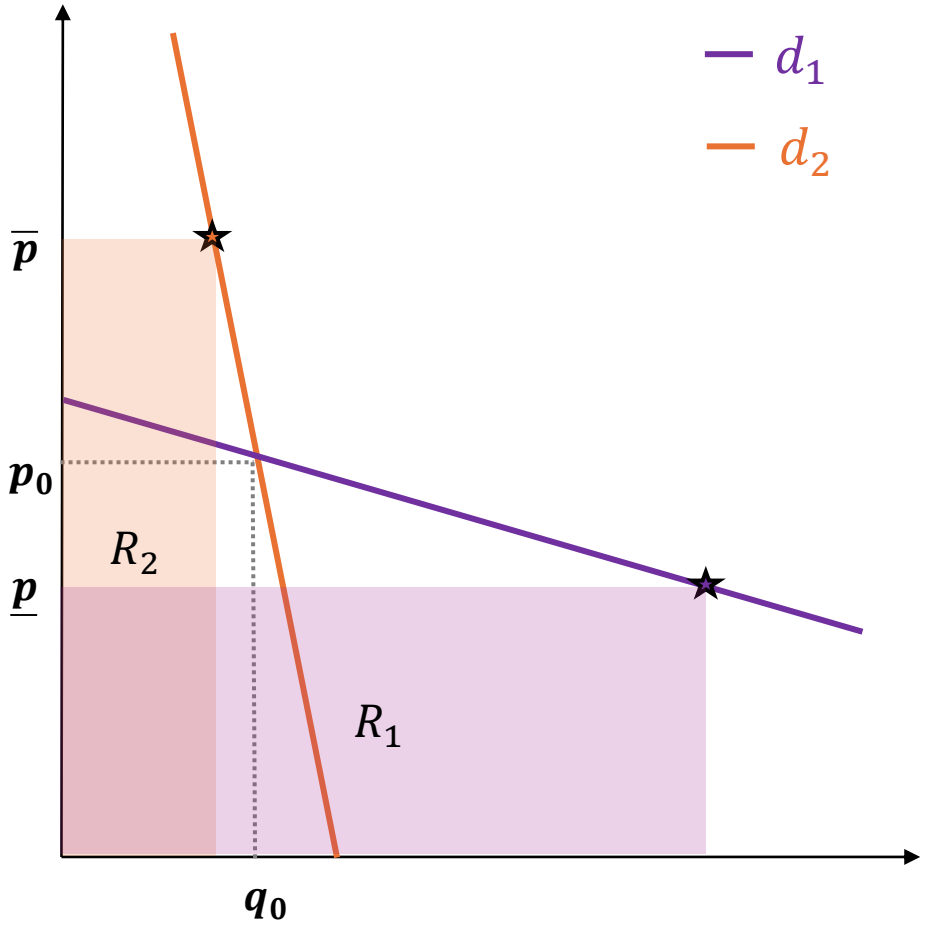
Locally informative → prices always rise

Applications: binary state space

Not downward closed

Not effectively downward closed

So myopic policy not optimal



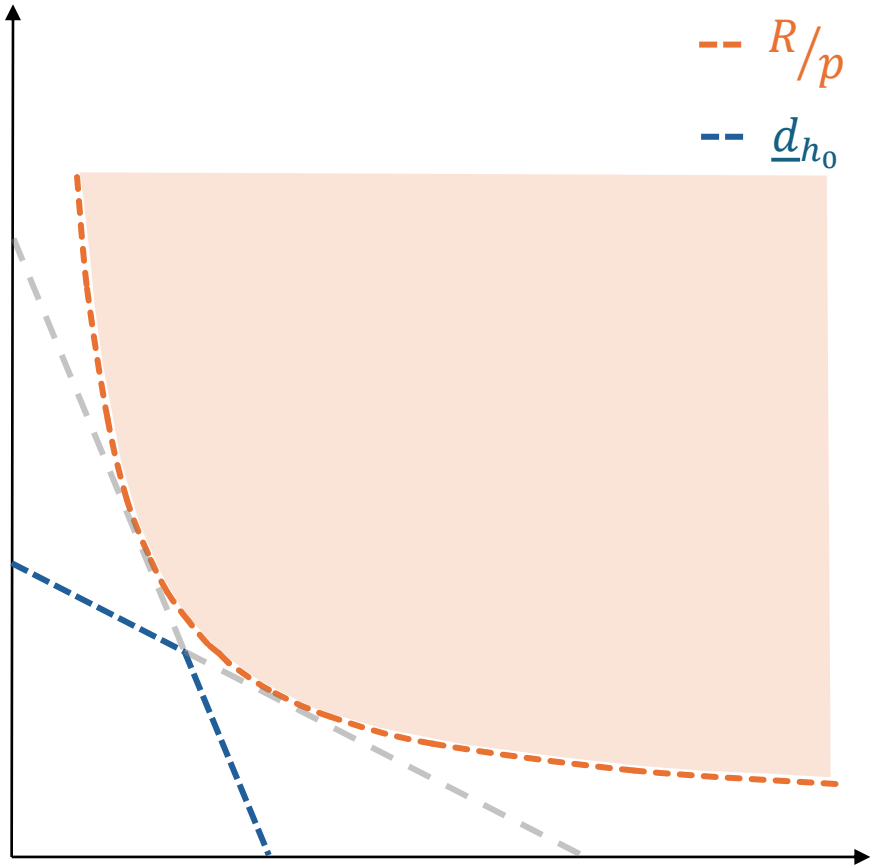
Applications: known market size

Let $R > 0$ and $\underline{\tau}, \overline{\tau} < 0$

Assumption 4: Ω is set of demand curves $d: \mathbb{R}_+ \to \mathbb{R}_+$ where

1. $\max_p p \cdot d(p) \geq R$,
2. $\underline{\tau} \leq \frac{d(x)-d(y)}{x-y} \leq \overline{\tau}$ for all $x \neq y$ where $d(x), d(y) > 0$.

Applications: known market size



Not effectively downward closed → optimal strategy is not myopic

Outline

1. Model
2. Special case
3. General case
4. Discussion

Price experimentation literature

Long-run learning:
Rothschild (1974), Easley and Kiefer (1988), Aghion et al (1991)…

Binary state space:
McLennan (1984), Rustichini and Wolinsky (1994), Keller and Rady (1999)…

One or Two-period models:
See references in Mirman et al (1993)…

Simple heuristics or asymptotic objectives:
Lobo and Boyd (2003), Besbes and Zeevi (2009), Cho and Libgober (2025)…

Ambiguity aversion and pricing

Ilut, Valchev, and Vincent (2020)	**Malladi (2025)**
worst case over convex set of full-support priors	worst case over states rather than priors
kinked demand	kinked demand
firm revisits past prices (price stickiness)	firm never revisits past prices *passive case*: prices rise *active case*: slows price discovery

Open questions on price experimentation

1. Single buyer and bid shading (ratchet effect)
2. Storable goods
   - For dynamic pricing with durable goods and robustness to buyer learning, see: Li, Libgober and Mu (2025)
3. Multiple goods (substitutes vs complements?)
4. Externalities (e.g., toll pricing)
5. Adverse selection (e.g., insurance deductibles)
6. Other criteria like minmax regret

Summary

Start with credible assumptions about what monopolist knows

Derive policies that maximize profit guarantees

If state space is sufficiently rich: increasing prices and profits

Search and Learning

Banchio and Malladi (**2024**) and **Malladi** (**2023**)

New elements here

Flexibility: Show how problem changes with Π and Ω

Tractability: Solve problems where optimal strategy is not myopic

Connect back to themes from Alessandro's and Rahul's talks:

Dealing with **non-uniqueness**

Justifying **alternative objectives**

Terminology

Experimentation: unknown flow payoff, no stopping

Search: known flow payoff (search cost), terminal reward upon stopping

Robust search may capture discovery process

- testing dosage of a new compound in a drug
- buying your first tennis racquet

How does this process depend on what we know *a priori*?

Rediscovery vs original discovery

How do agents search when they **know that good outcomes exist** but don't know where to find them?

Ex 1: Solving a homework problem

- Teacher guarantees a solution exists
- Student tries to rediscover it by trial and error

Ex 2: Catching up to an innovative competitor

- Competitor keeps recipe as a trade secret
- Firm does R&D to rediscover it

Rediscovery is simpler than original discovery

It's really easy to copy something you know works. One of the reasons people don't talk about why it's so easy is that ***you have the conviction to know that it's possible****. And so after a research lab does something, even if you don't know exactly how they did it, it's --- I won't say easy, but --- doable to go off and copy it [...] What is really hard [...] is the repeated ability to go off and do something new and totally unproven... A lot of organizations talk about the ability to do this. There are very few that do, across any field.*

Sam Altman on OpenAI's competitors

Behavioral theory of the firm

Problemistic search (Cyert and March 1963; Simon 1962)

"Search within the firm is problem-oriented…innovation by a competitor"

Rugged landscapes (Levinthal 1997; Rivkin 2000; Callander 2011)

Mapping from choices to performance is complex and unpredictable

Rediscovery: optimal "problemistic search on rugged landscapes"

Outline

1. Model
2. Simple search policies
3. Proof ideas
4. Discussion

Rugged landscapes and rediscovery

Fix $L, k \in \mathbb{R}_+$

Q is the set of all quality indices that:

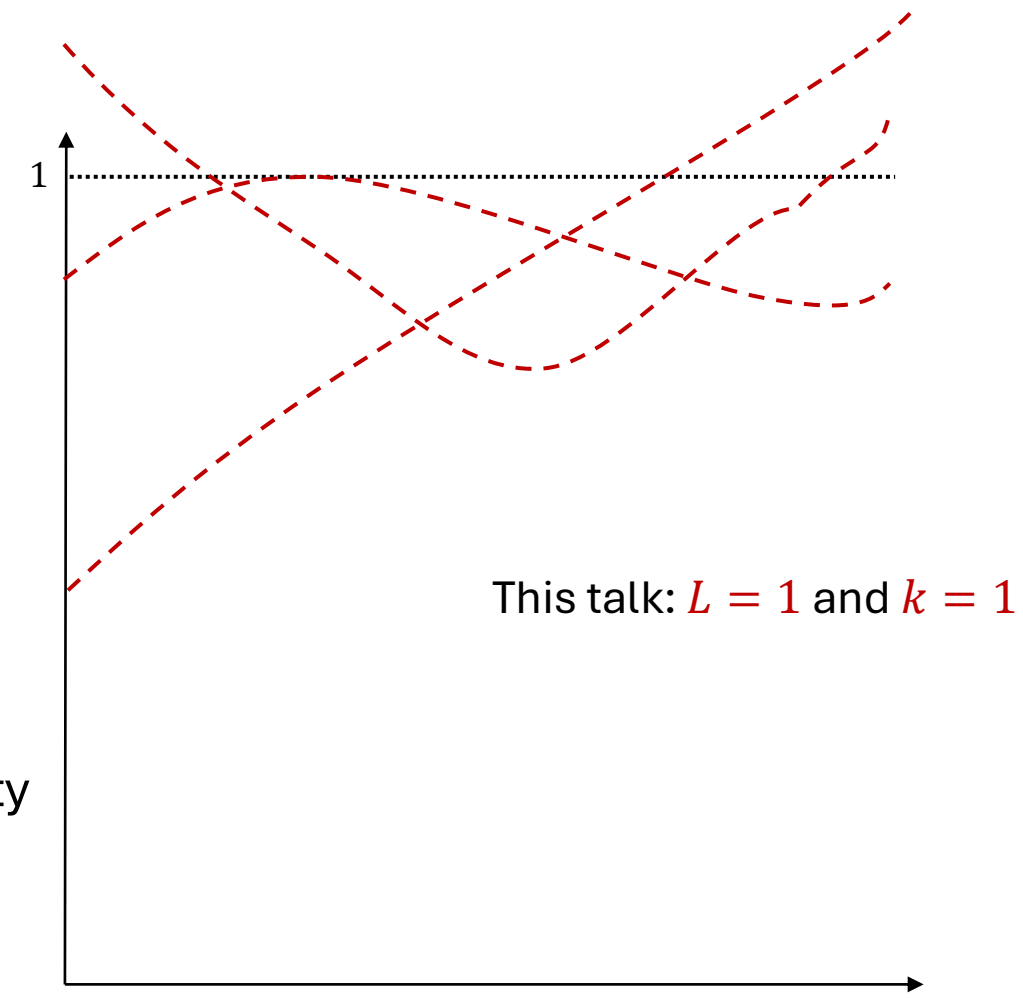
1. are L–Lipschitz continuous
2. attain value k somewhere

item's location in S: observable attribute

continuity: observably similar = similar quality

L : prior on ruggedness

k: target



This talk: $L = 1$ and $k = 1$

Payoff and objective

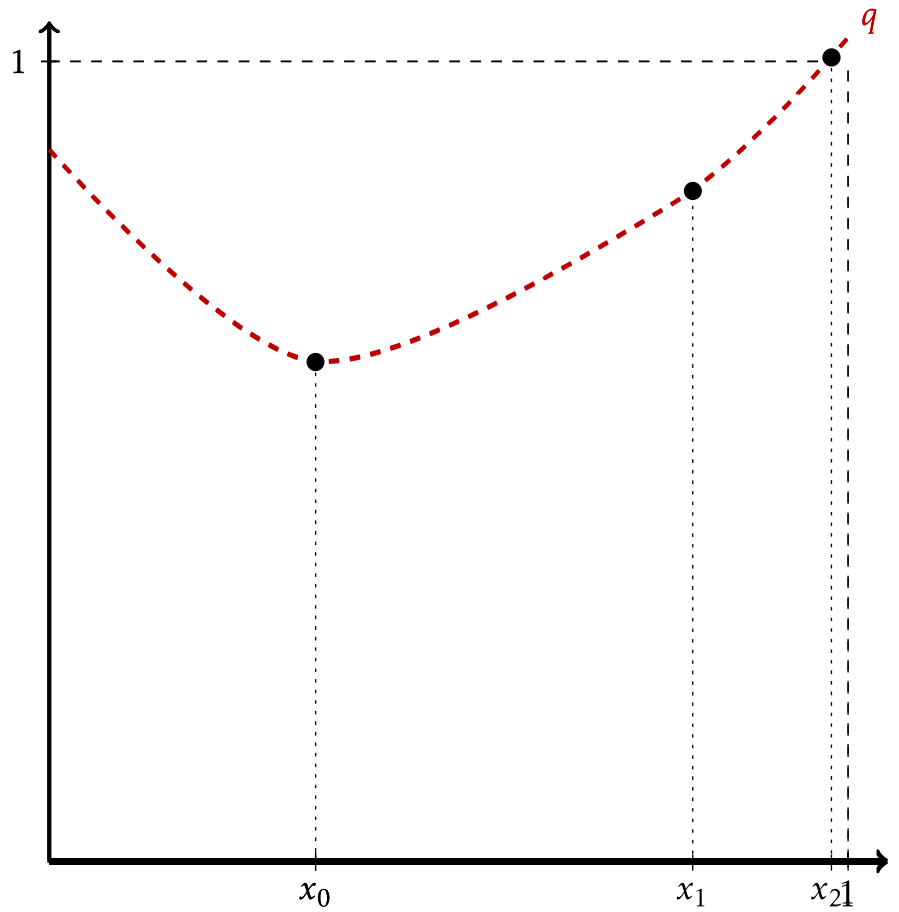
Payoff to stopping at h_0 is 0

Payoff to stopping at $h_t \in H$:

$$p(h_t) = \max_{i \in \{0,\ldots,t-1\}} z_i - c \cdot t$$

σ^* is (ex ante) **optimal** if it maximizes guaranteed payoff, i.e., at $h = h_0$:

$$\sigma^* \in \operatorname{argmax}_{\sigma \in \Sigma} \min_{q \in Q} p\left(h_q^\sigma\right)$$



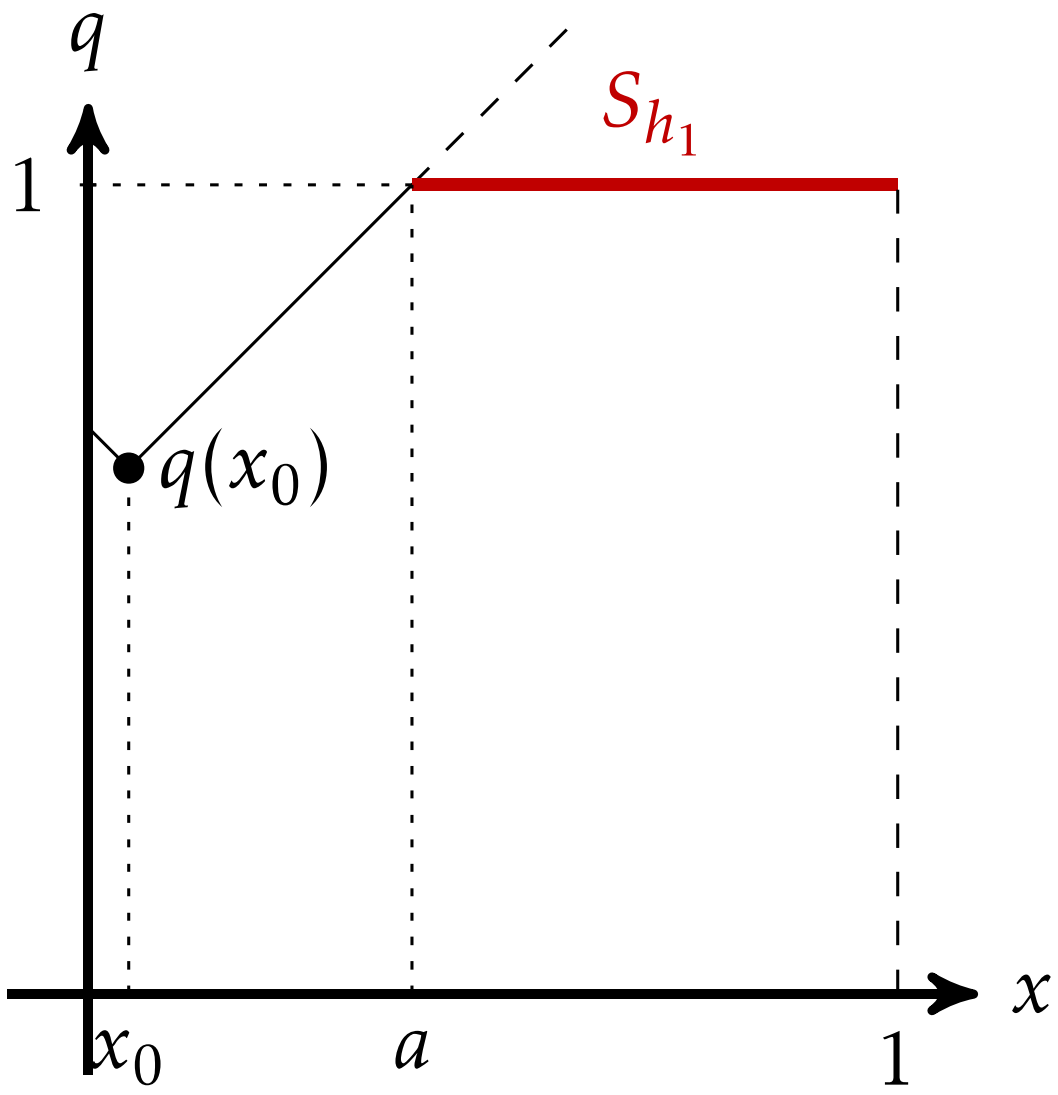
Outline

1. Model
2. Simple search policies
3. Proof ideas
4. Discussion

Search window

Region of S where target may be found:

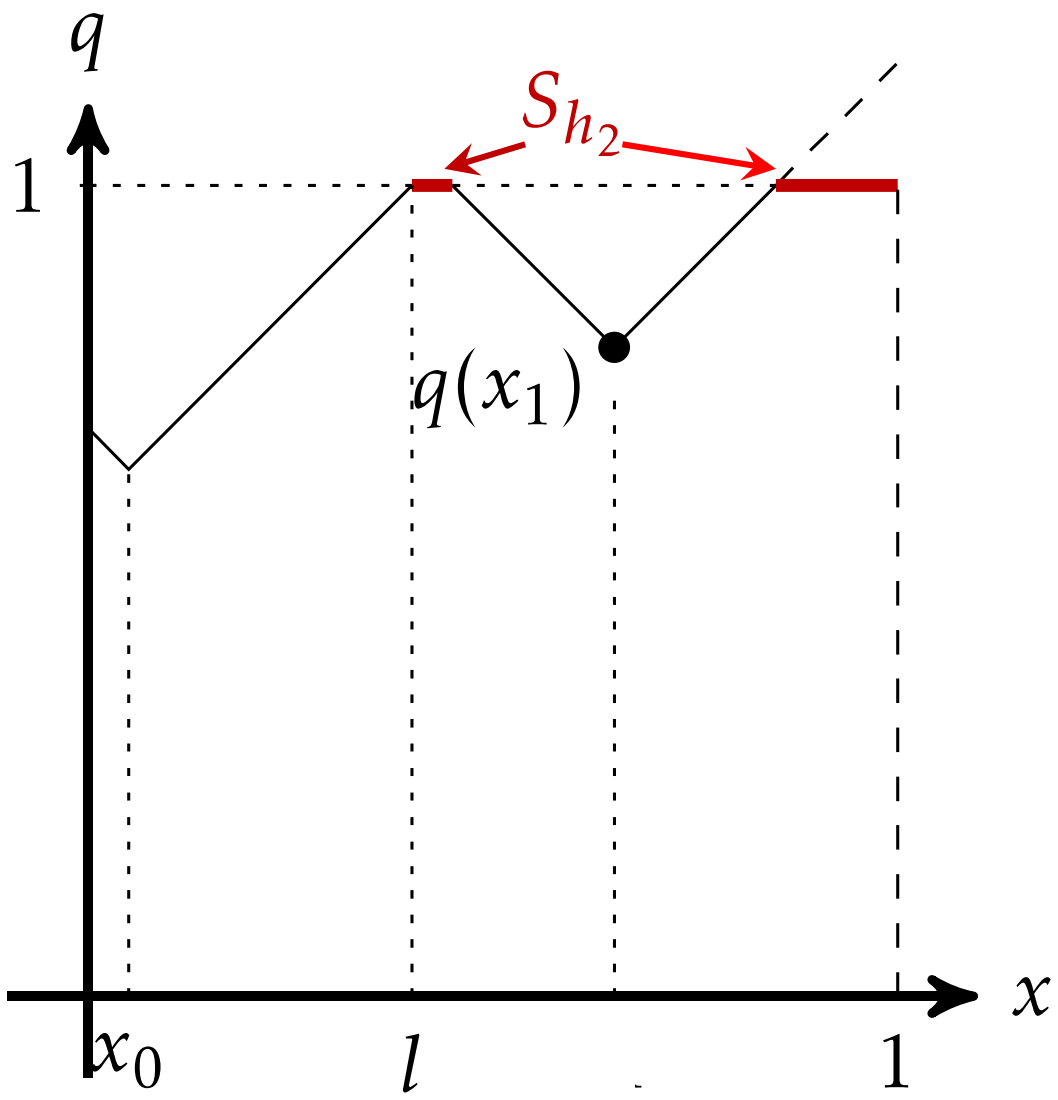
$S_h \equiv \{x \in S | \exists q \in Q_h \; s.t. \; q(x) = 1\}$



Search window

Region of S where target may be found:

$S_h \equiv \{x \in S | \exists q \in Q_h \; s.t. \; q(x) = 1\}$



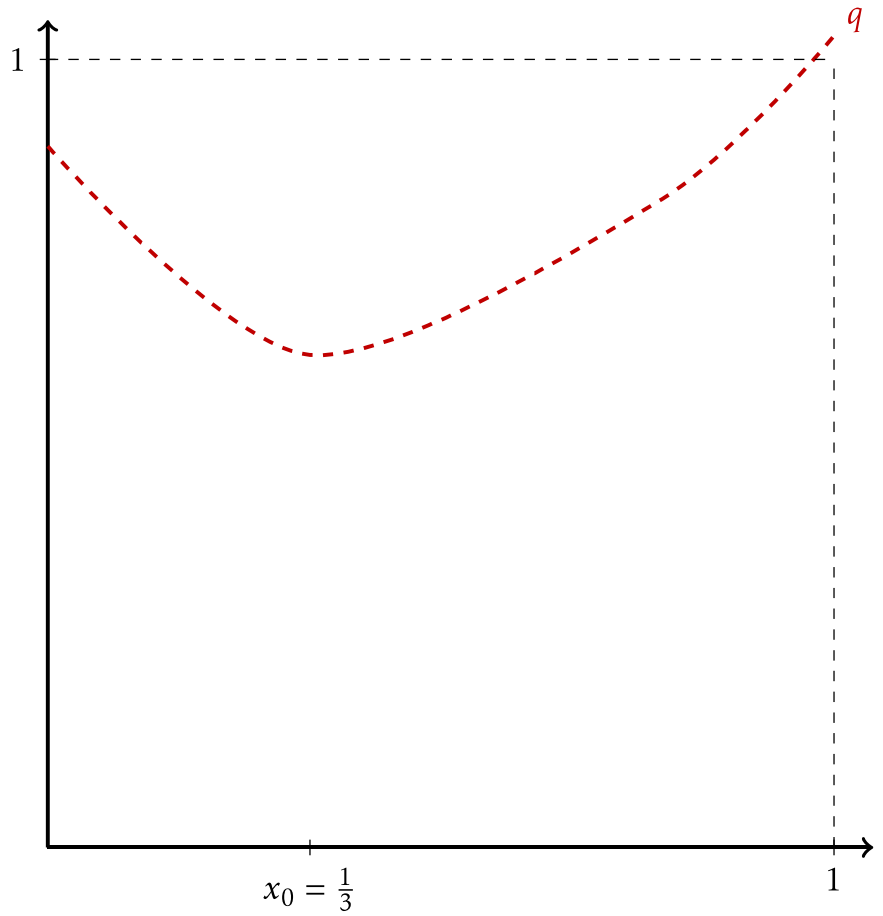
Example search policy

$\lambda(h)$: length of search window at h

At every reachable history $h \in H^{\sigma}$

→ stop if $\max_{i \in \{0,\ldots,t-1\}} z_i \geq 1 - \frac{1}{3}\lambda(h_t)$

→ search at $1 - \frac{2}{3}\lambda(h_t)$



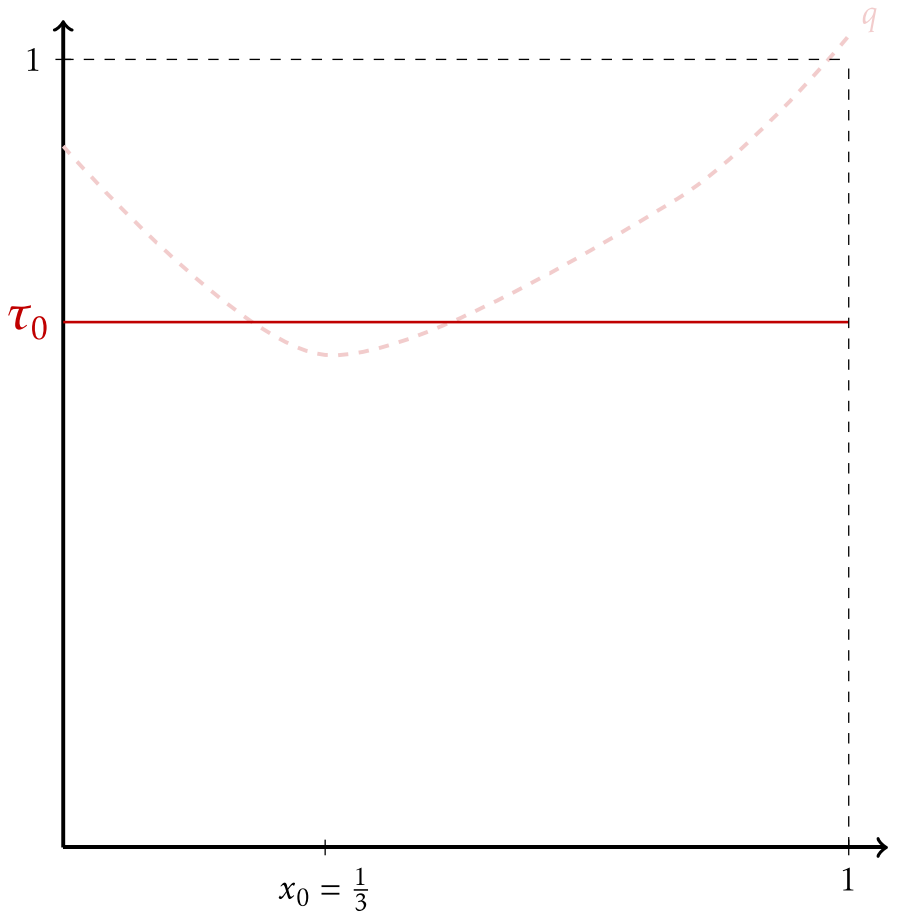
Example search policy

$\lambda(h)$: length of search window at h

At every reachable history $h \in H^\sigma$

→ stop if $\max_{i \in \{0,\ldots,t-1\}} z_i \geq 1 - \frac{1}{3}\lambda(h_t)$

→ search at $1 - \frac{2}{3}\lambda(h_t)$



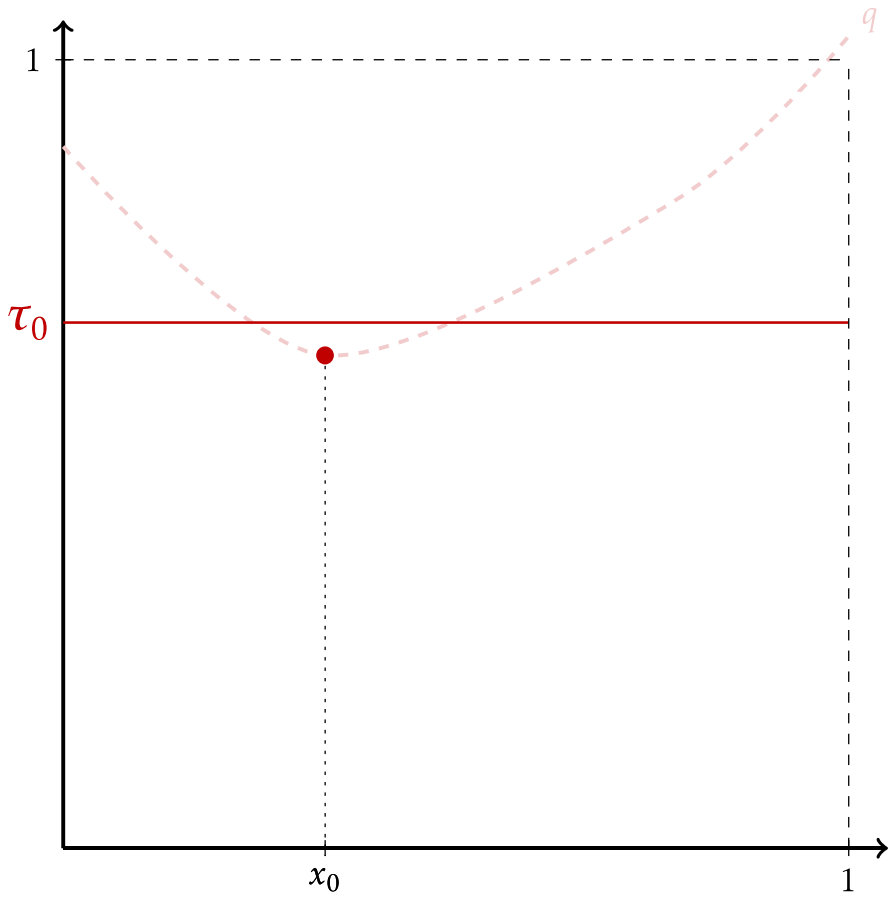
Example search policy

$\lambda(h)$: length of search window at h

At every reachable history $h \in H^\sigma$

→ stop if $\max_{i \in \{0,\ldots,t-1\}} z_i \geq 1 - \frac{1}{3}\lambda(h_t)$

→ search at $1 - \frac{2}{3}\lambda(h_t)$



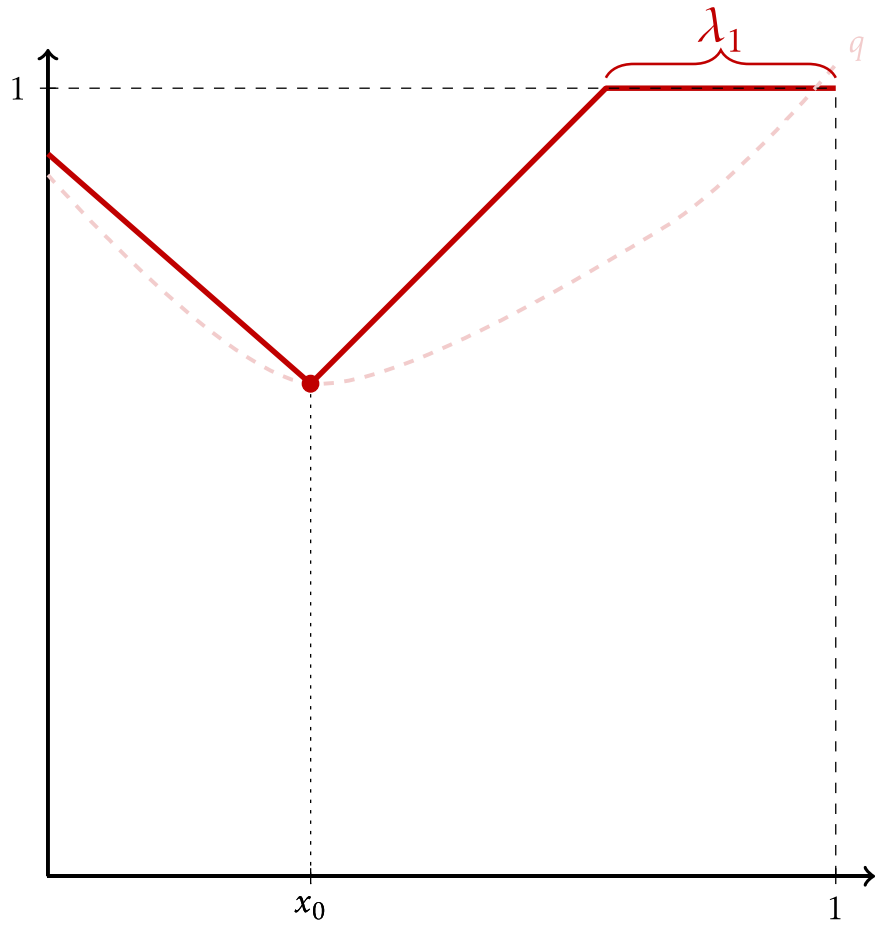
Example search policy

$\lambda(h)$: length of search window at h

At every reachable history $h \in H^{\sigma}$

→ stop if $\max_{i \in \{0,\ldots,t-1\}} z_i \geq 1 - \frac{1}{3}\lambda(h_t)$

→ search at $1 - \frac{2}{3}\lambda(h_t)$



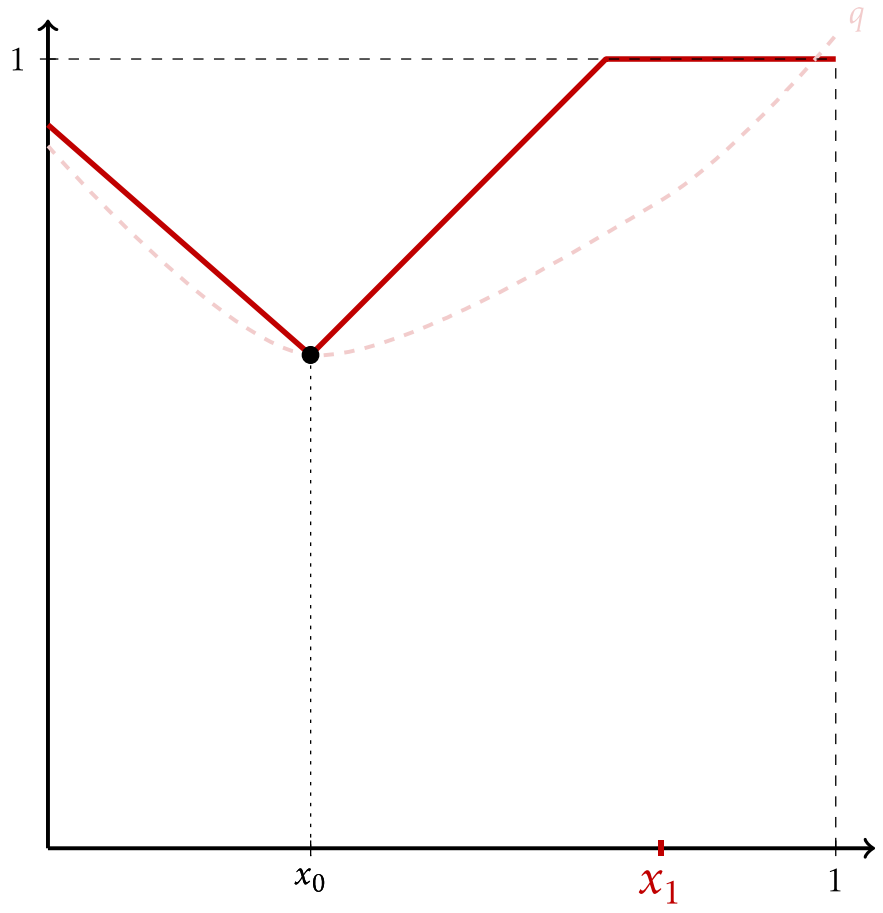
Example search policy

$\lambda(h)$: length of search window at h

At every reachable history $h \in H^{\sigma}$

→ stop if $\max_{i \in \{0,\ldots,t-1\}} z_i \geq 1 - \frac{1}{3}\lambda(h_t)$

→ search at $1 - \frac{2}{3}\lambda(h_t)$



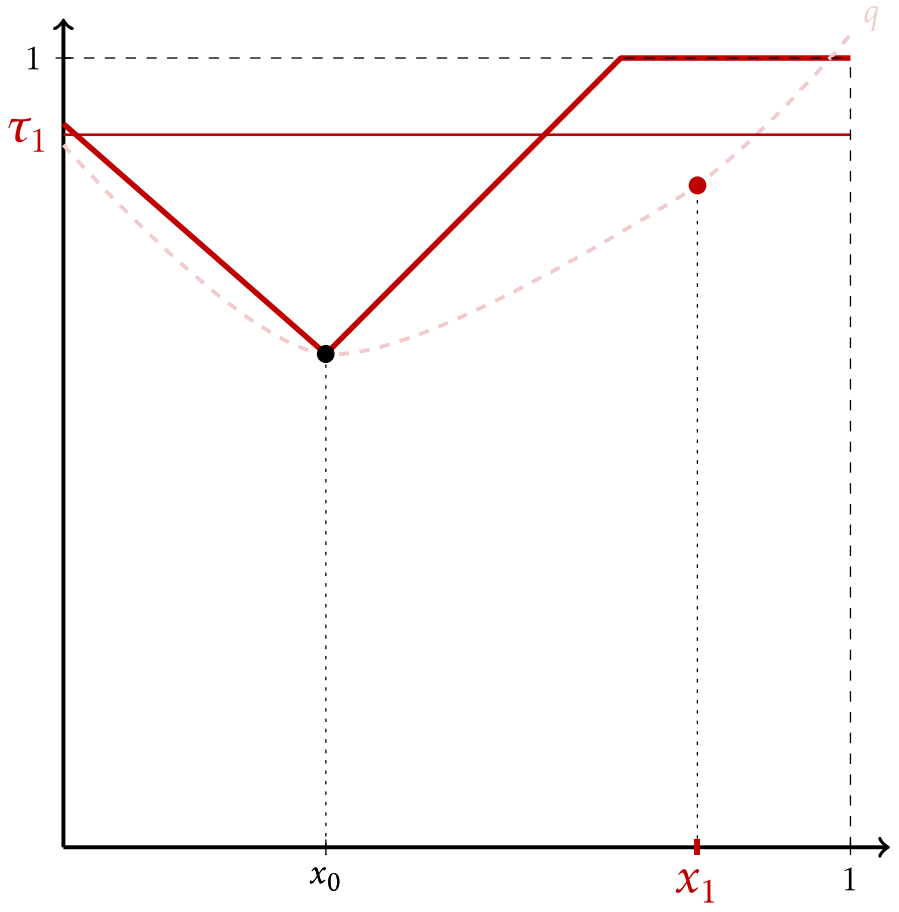
Example search policy

$\lambda(h)$: length of search window at h

At every reachable history $h \in H^{\sigma}$

→ stop if $\max_{i \in \{0,\ldots,t-1\}} z_i \geq 1 - \frac{1}{3}\lambda(h_t)$

→ search at $1 - \frac{2}{3}\lambda(h_t)$



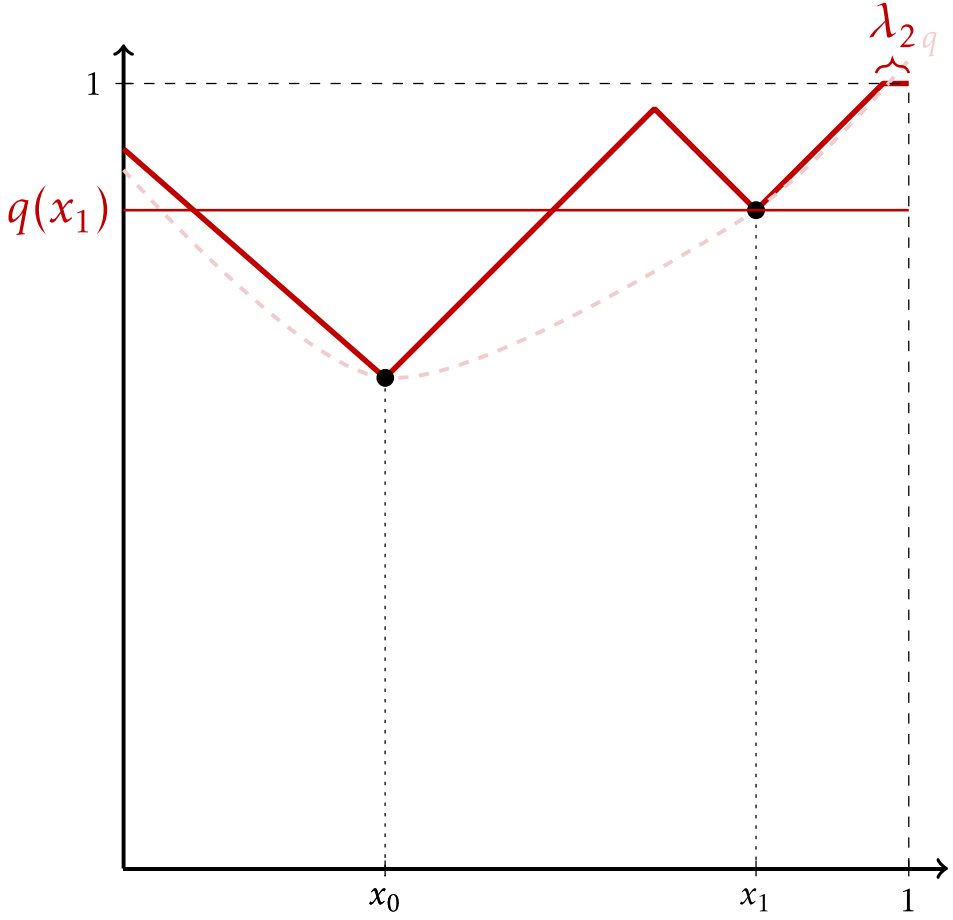
Example search policy

$\lambda(h)$: length of search window at h

At every reachable history $h \in H^{\sigma}$

→ stop if $\max_{i \in \{0,\ldots,t-1\}} z_i \geq 1 - \frac{1}{3}\lambda(h_t)$

→ search at $1 - \frac{2}{3}\lambda(h_t)$



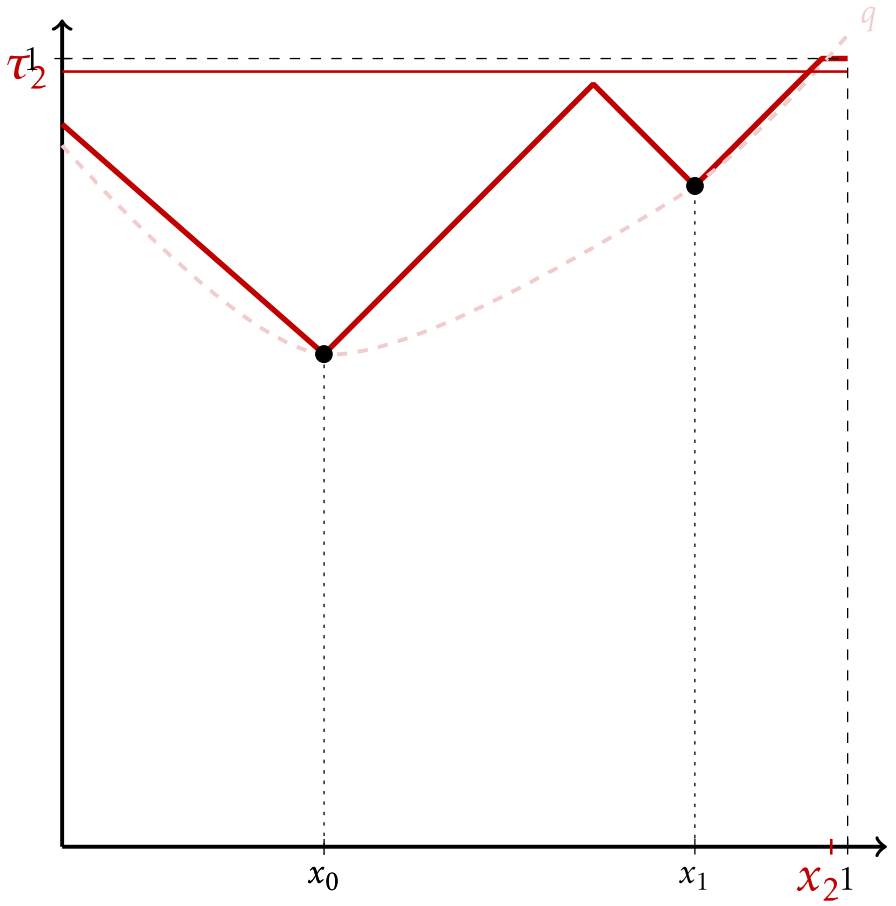
Example search policy

$\lambda(h)$: length of search window at h

At every reachable history $h \in H^{\sigma}$

→ stop if $\max_{i \in \{0,\ldots,t-1\}} z_i \geq 1 - \frac{1}{3}\lambda(h_t)$

→ search at $1 - \frac{2}{3}\lambda(h_t)$



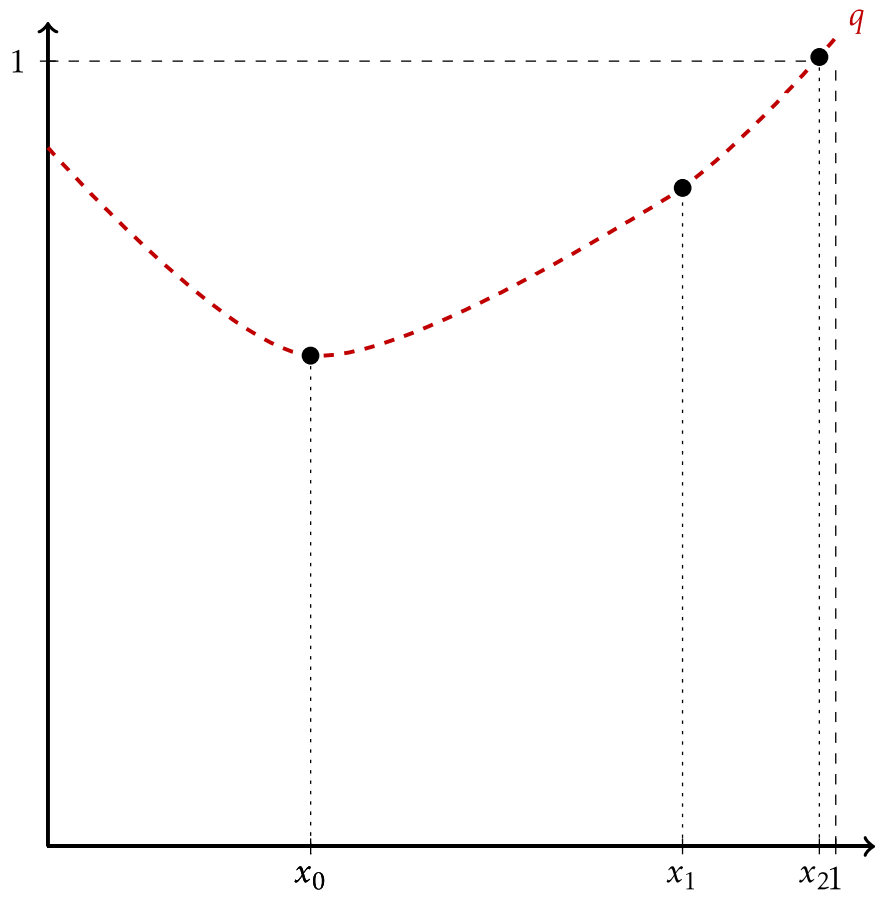
Example search policy

$\lambda(h)$: length of search window at h

At every reachable history $h \in H^{\sigma}$

→ stop if $\max_{i \in \{0,\ldots,t-1\}} z_i \geq 1 - \frac{1}{3}\lambda(h_t)$

→ search at $1 - \frac{2}{3}\lambda(h_t)$



Example search policy

Search policy σ

- explores **directionally**
- is an **index policy**: depends only on search window length
- stops only if quality exceeds some **threshold**
- **does not invoke recall**

Main result

Theorem: There exists a sequentially optimal policy that is directional, threshold, index, does not invoke recall.

Rediscovery is a simple process

1. Searcher can explore the space freely
   Suffices to **search in increments**
2. Searcher can use complex stopping rules
   Threshold rules are optimal
3. Searcher has perfect recall
   Recall is never invoked
4. Histories are complex
   Search window length is **only state variable**

None of these properties holds in our model of original discovery

Outline

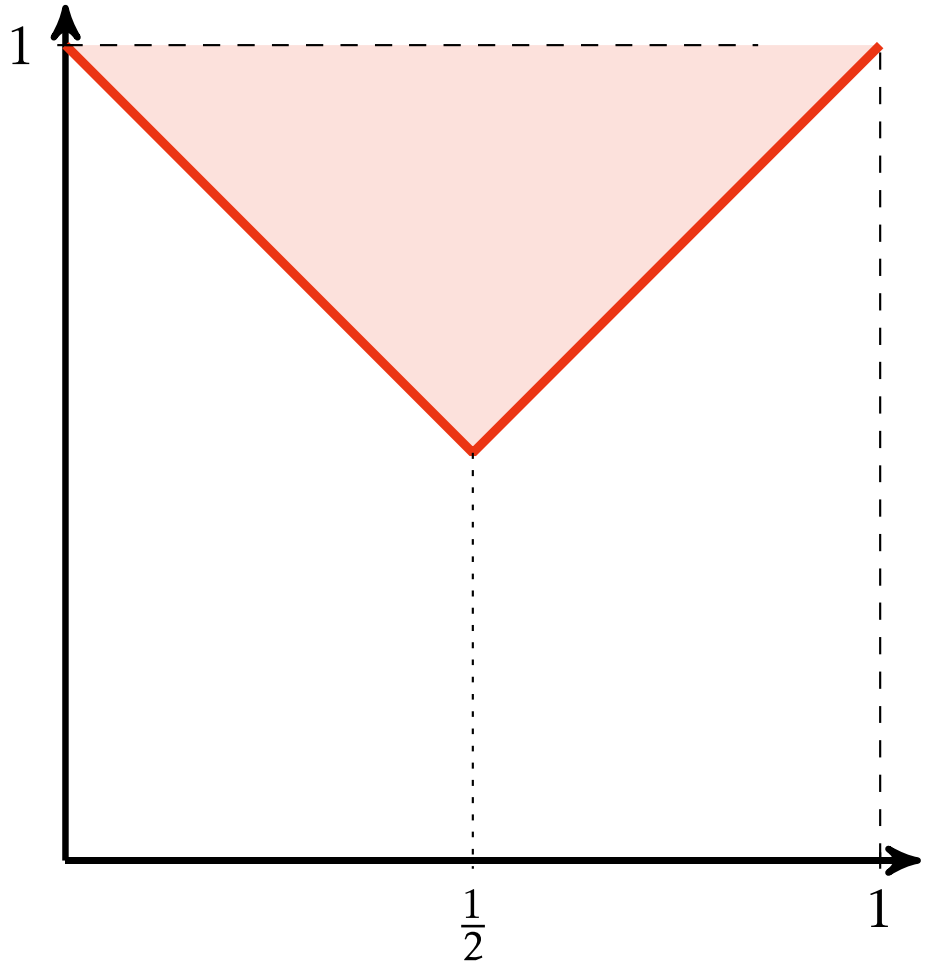
1. Model
2. Simple search policies
3. Proof ideas
4. Discussion

Two period case

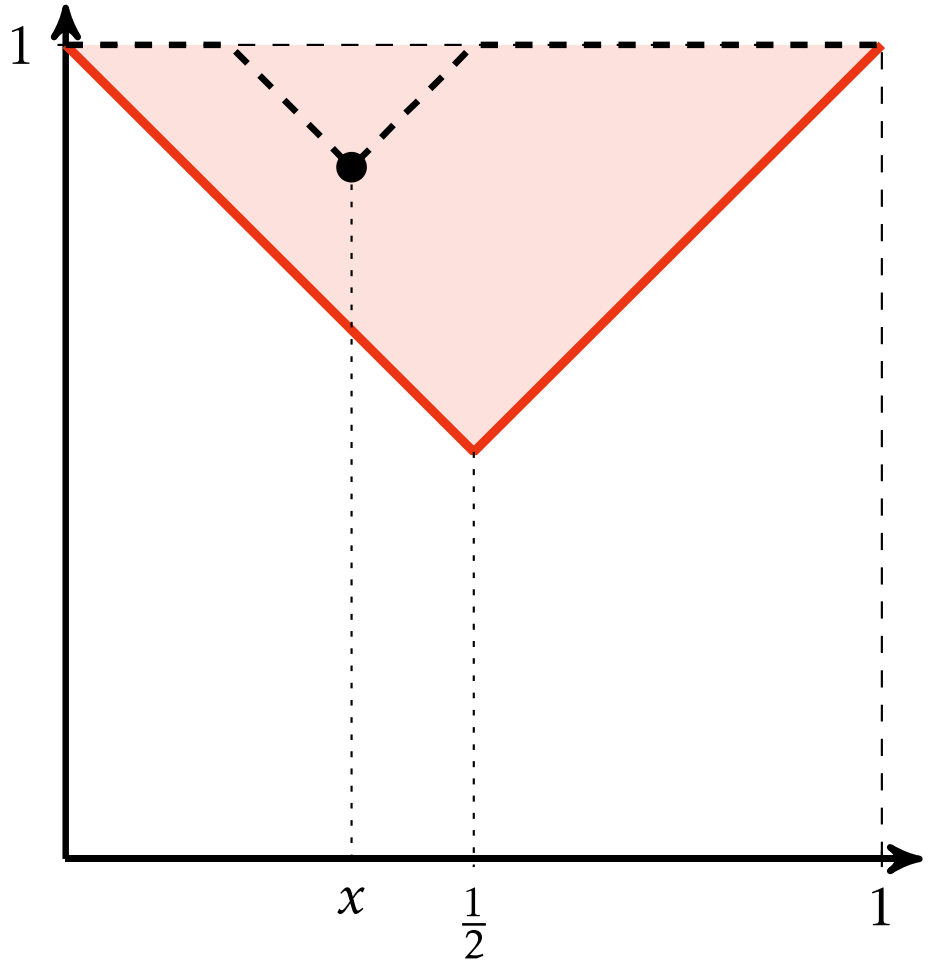
Assumption: suppose search must end in two periods

Searcher’s optimal strategy hedges against two ‘risks’

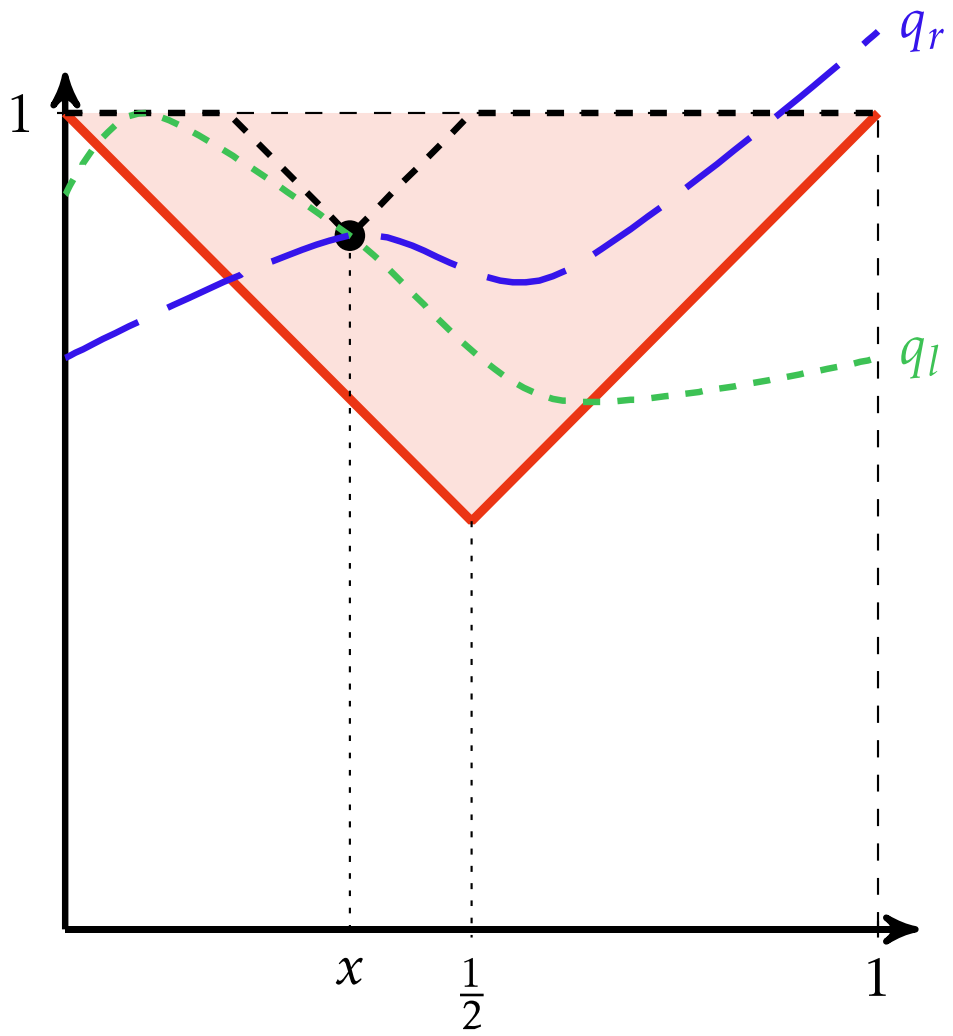
Bifurcation risk



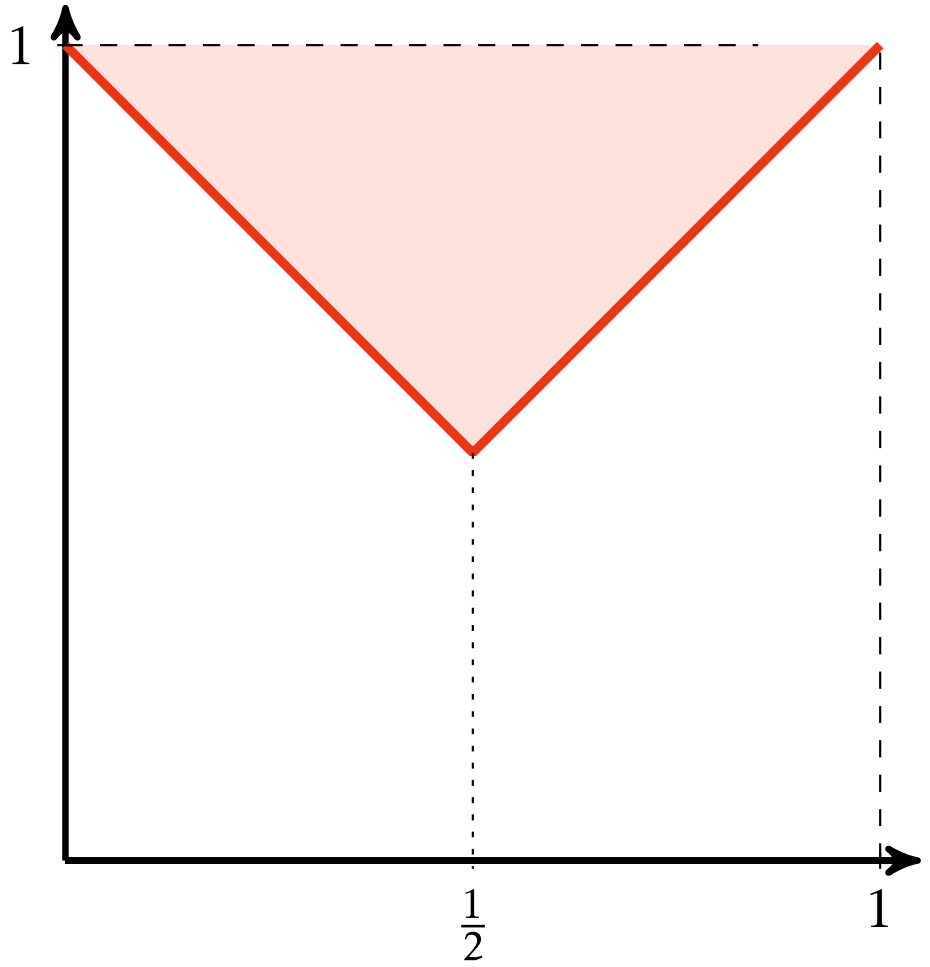
Bifurcation risk



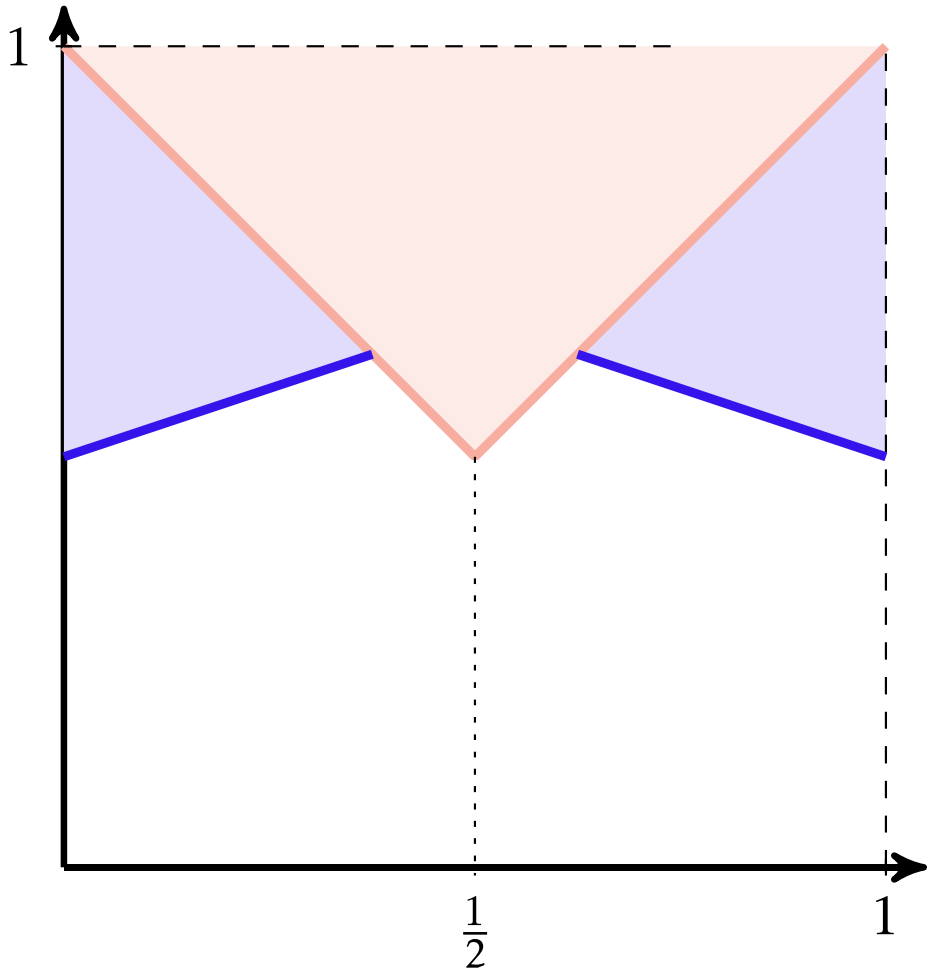
Bifurcation risk



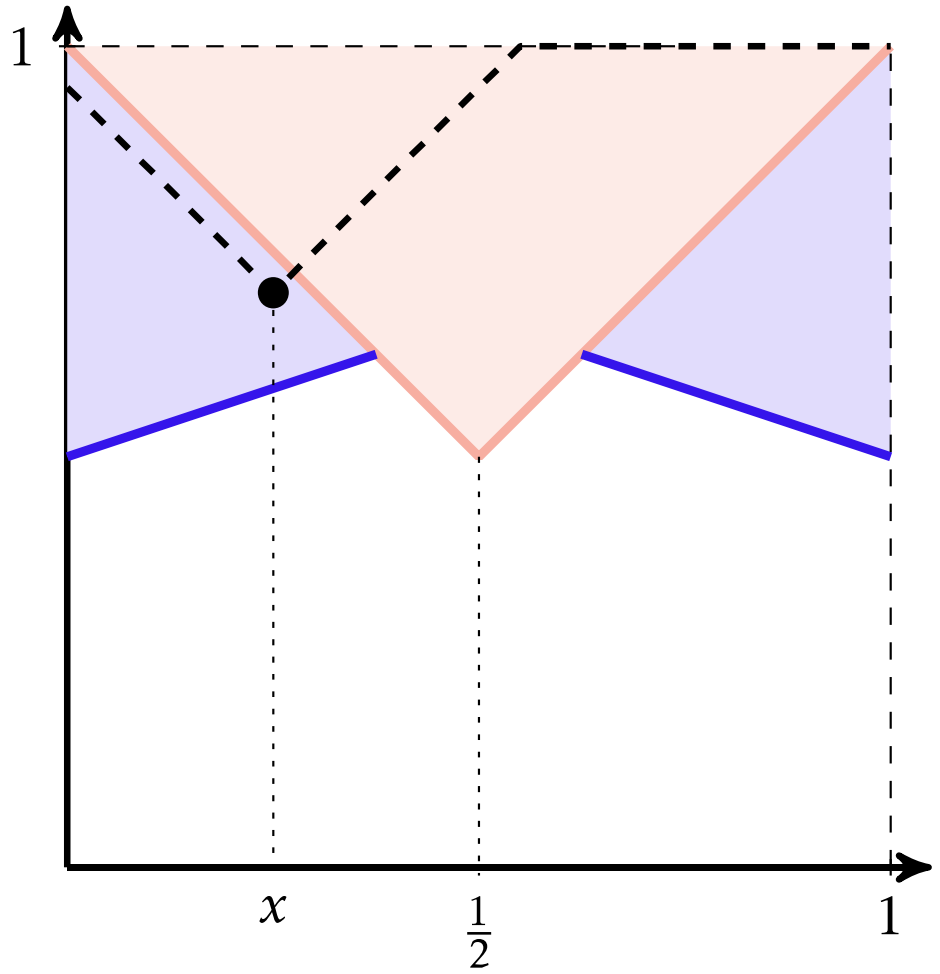
Bifurcation risk



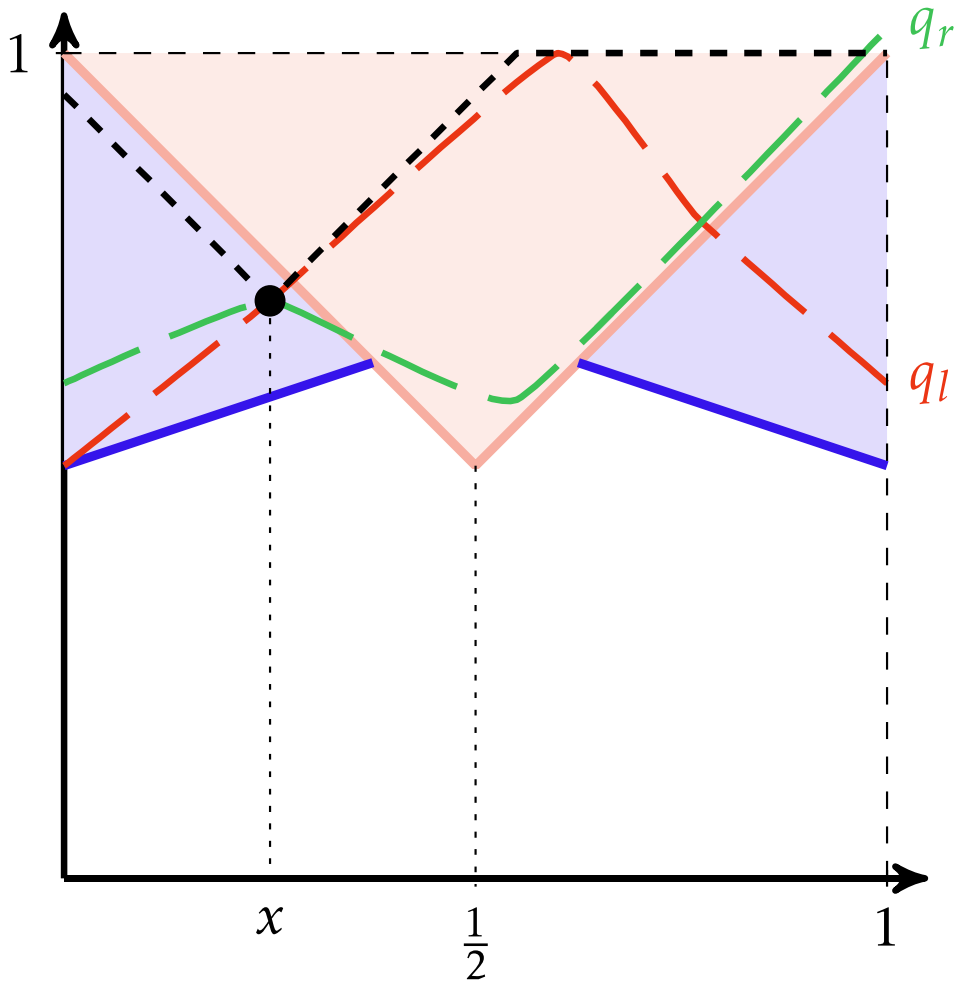
Directional risk



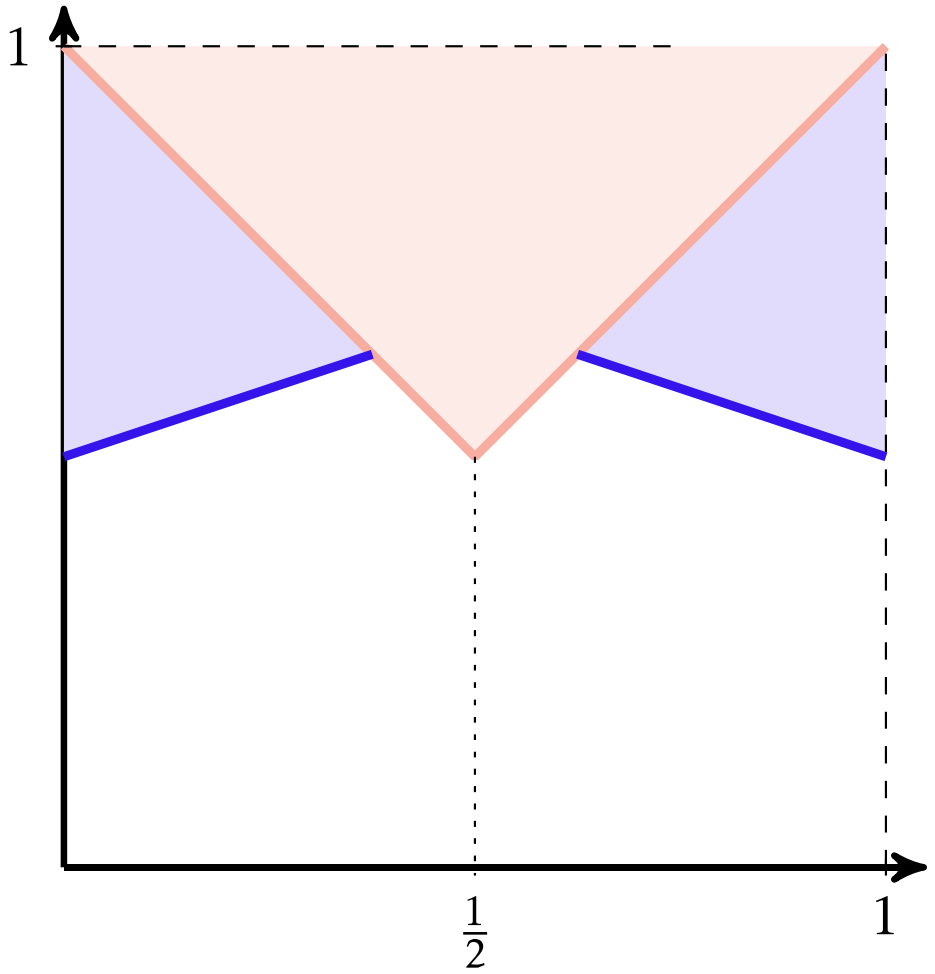
Directional risk



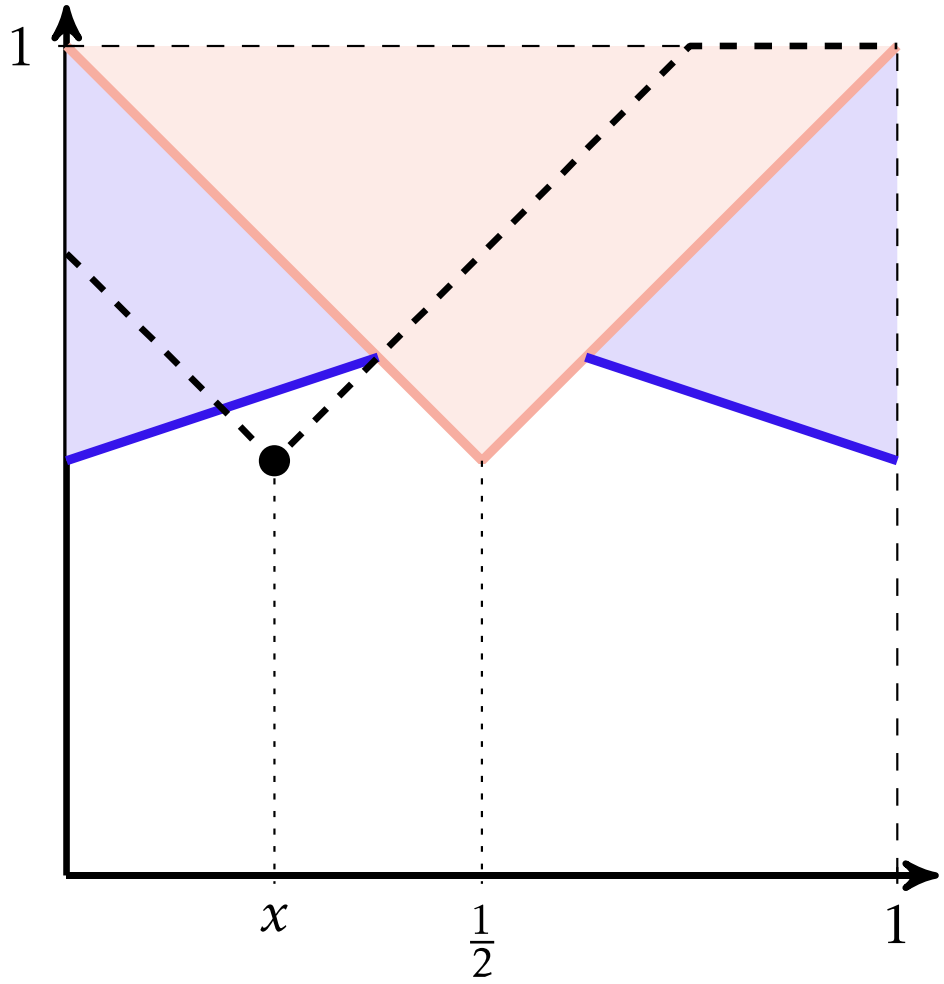
Directional risk



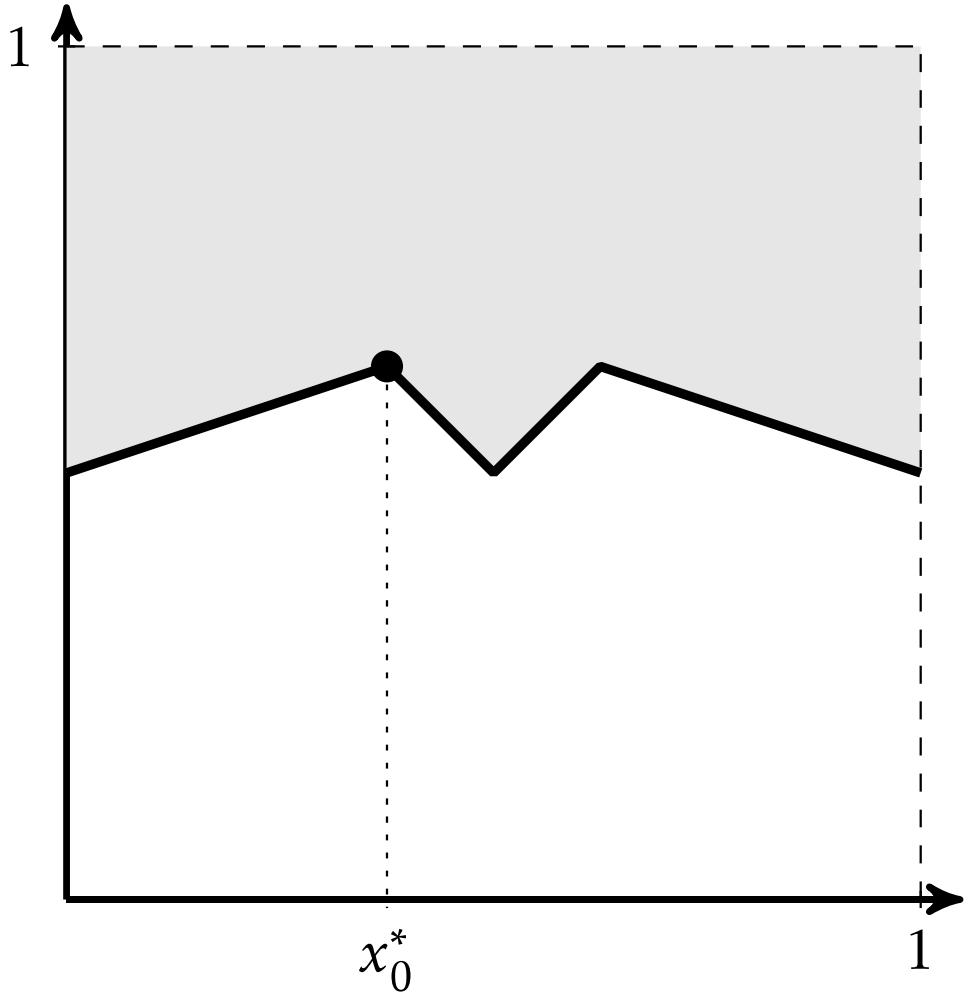
Directional risk



Continuation region



Optimal search policy



Outline

1. Model
2. Simple search policies
3. Proof ideas
4. Discussion

Uniqueness

Sequentially optimal strategy is not unique

- e.g., left to right or right to left
- could be many more

Selection: distance travel costs

- e.g., overhauling design more expensive than tinkering with it
- picks a directional strategy

Robustness: what holds true across all seq opt strategies?

- No-recall and threshold
- Holds true even in more general search + state spaces

Rediscovery

Agent benchmarks against what was **previously discovered**

$$\min_{\sigma \in \Sigma} \max_{q \in Q} k - p\left(h_q^{\sigma}\right)$$

where Q is all L-Lipschitz functions on [0, 1] **that attain a value of** k

Original discovery

Agent benchmarks against what is **discoverable**

$$\min_{\sigma \in \Sigma} \max_{q \in Q} k_q - p\left(h_q^{\sigma}\right)$$

where Q is all L-Lipschitz functions on [0, 1] and $k_q \equiv \max_{x \in [0,1]} q(x)$

Seq. opt. strategies not directional, threshold, index, or no-recall

Aspects of rediscovery but not original discovery

1. **Process of elimination**:

   Bad discovery → remove similar choices from consideration

2. **Satisficing**

   "[so] long as the problem is not solved, search will continue." (Cyert and March 1963)

3. **No discouragement effects (increasing threshold)**

   Emboldened by failure because objective is attainable.

Alternative robust objectives

Many objections to using maxmin

Sometimes, maxmin seems okay, but not regret or competitive ratio

- Why should firm care about what is achievable?
- Which functional form is right?
- "CS not econ"

One solution: work with a general functional form

- Malladi (2023): payoff is $U(q(x), k_q)$, increasing in first, decreasing in second
- Objective is just maxmin with reference-dependent utility
- Economists have many examples for such preferences

More questions

1. Strategic search/experimentation
2. Motivating agents to search
3. More general information acquisition models
4. Misspecification? (see Lars Hansen's work)

Dynamic Consistency

Malladi (2023)

Robustness and dynamics

Carroll (**2019**): "*Trying to write dynamic models with non-Bayesian decision makers leads to well-known problems of dynamic inconsistency except in special cases (e.g. Epstein & Schneider 2003). This may be one reason why there has been relatively little work so far on robust mechanism design in dynamic settings.*"

Without DC: full commitment? consistent planning? naivety?

e.g., see Auster, Che, Mierendorff (2024)

Why do sequentially optimal strategies exist in our setting?

Weak dynamic consistency

Consider strategy $\sigma \in \Sigma$ and history $h \in H$

These determine search path and payoff for any $q \in \Omega_h$

Extend preferences to all acts $f: \Omega \to \mathbb{R}$ in $\mathcal{A}$

For any $f, g \in \mathcal{A}$ and $h \in H$,

$$f \succsim_h g \iff \inf_{q \in \Omega_h} f(q) \geq \inf_{q \in \Omega_h} g(q)$$

Weak dynamic consistency

Proposition: For all $h \in H$ and any unsearched item $x \in S$, if $f \gtrsim_{h+(x,z)} g$ for all possible qualities z, then $f \gtrsim_h g$.

Proof: For all z

$$\inf_{q \in \Omega_{h+(x,z)}} f(q) \geq \inf_{q \in \Omega_{h+(x,z)}} g(q)$$

and,

$$\Omega_h = \bigcup_z \Omega_{h+(x,z)}$$

so,

$$\inf_{q \in \Omega_h} f(q) \geq \inf_{q \in \Omega_h} g(q)$$

Existence of sequentially optimal strategies

Take an ex ante optimal strategy σ

Modify σ to improve the payoffs at some future history

By the proposition, σ remains ex ante optimal

Intuition:
if that history was off path, remains so after modification
if on path, nature gives weakly better outcome than before

Conclusion

Robustness and learning

Actions DMs take can affect what they learn

Learning can be important in settings w/ maxmin preferences

- Tractable too!

Robust mech design with limited commitment:

Li, Libgober and Mu (2025), Koh and Sanguanmoo (2025)…

More on ambiguity aversion and learning:

Survey: Ilut and Schneider (2022)