Dynamic Delegation of Experimentation

Yingni Guo
Northwestern University

Delegating Experimentation

- I study how to manage innovation in a hierarchical organization.
- A principal delegates experimentation to an agent.

Delegating Experimentation

- I study how to manage innovation in a hierarchical organization.
- A principal delegates experimentation to an agent.

- **Private information**: the agent knows more precisely the prospect of the experimentation.
- **Misaligned preferences**: the agent prefers more experimentation.

Delegating Experimentation

- I study how to manage innovation in a hierarchical organization.
- A principal delegates experimentation to an agent.

- **Private information**: the agent knows more precisely the prospect of the experimentation.
- **Misaligned preferences**: the agent prefers more experimentation.
- **Tradeoff**: using the agent's information and constraining his bias.

What is the Optimal Delegation Contract?

- The principal has full commitment.
- The principal cannot use transfers.
- The principal can only impose limitations on the agent's behavior.

What is the Optimal Delegation Contract?

- The principal has full commitment.
- The principal cannot use transfers.
- The principal can only impose limitations on the agent's behavior.

- As new information arrives over time, how should the principal adjust the flexibility that the agent has?
- Is the optimal delegation contract time-consistent?

Short Answer: Main Result

- The optimal contract is **a cutoff rule** in the belief space and can be implemented as a sliding deadline:
  - The principal initially sets a deadline for experimentation;
  - Whenever encouraging information arrives, the deadline is extended;
  - The agent has full flexibility before the deadline but none after.

Short Answer: Main Result

- The optimal contract is **a cutoff rule** in the belief space and can be implemented as a sliding deadline:
  - The principal initially sets a deadline for experimentation;
  - Whenever encouraging information arrives, the deadline is extended;
  - The agent has full flexibility before the deadline but none after.
- The cutoff rule is time-consistent.

Short Answer: Main Result

- The optimal contract is **a cutoff rule** in the belief space and can be implemented as a sliding deadline:
  - The principal initially sets a deadline for experimentation;
  - Whenever encouraging information arrives, the deadline is extended;
  - The agent has full flexibility before the deadline but none after.
- The cutoff rule is time-consistent.
- The most promising products are under-experimented whereas less promising ones are over-experimented.

Examples: Within Firms and Others

- In-house innovation.
- Market learning.
- Public good provision.
- Research grants and funding.

Outline

1. Model
2. Single-player benchmark
3. Characterizing the policy space
4. Main results
5. More general results

Players, Tasks and States

- Time $t \in [0, \infty)$ is continuous.
- Two risk-neutral players $i \in \{\alpha, \rho\}$: Agent (he) and Principal (she).
- One unit of a divisible resource per unit of time.
- Agent continually splits the resource between two tasks:
  - S: known (deterministic flow) payoff;
  - R: unknown **state** of world $\omega \in \{0, 1\}$.

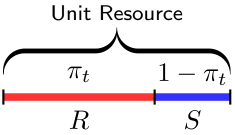
Players, Tasks and States

- Time $t \in [0, \infty)$ is continuous.
- Two risk-neutral players $i \in \{\alpha, \rho\}$: Agent (he) and Principal (she).
- One unit of a divisible resource per unit of time.
- Agent continually splits the resource between two tasks:
  - S: known (deterministic flow) payoff;
  - R: unknown **state** of world $\omega \in \{0, 1\}$.

For the talk: focus on Poisson bandits (conclusive news).
All results generalize to Lévy bandits.

Tasks and Payoffs

Over $[t, t + dt)$:



- S yields to player i

$$(1 - \pi_t)s_i dt.$$

- R yields a success with probability

$$\pi_t \lambda \omega dt.$$

- Each success is worth h_i to player i.

Tasks and Payoffs (cont.)

- Conditional on ω, the expected payoff increment to player i is

$$(1-\pi_t)s_i dt + \pi_t \lambda h_i \omega dt = \begin{bmatrix} s_i & \lambda h_i \omega \end{bmatrix} \cdot \begin{bmatrix} (1-\pi_t)dt \\ \pi_t dt \end{bmatrix}.$$

Tasks and Payoffs (cont.)

- Conditional on ω, the expected payoff increment to player i is

$$(1-\pi_t)s_i dt + \pi_t \lambda h_i \omega dt = \boxed{\begin{bmatrix} s_i & \lambda h_i \omega \end{bmatrix}} \cdot \begin{bmatrix} (1-\pi_t)dt \\ \pi_t dt \end{bmatrix}.$$

Tasks and Payoffs (cont.)

- Conditional on ω, the expected payoff increment to player i is

$$(1-\pi_t)s_i dt + \pi_t \lambda h_i \omega dt = \boxed{\begin{bmatrix} s_i & \lambda h_i \omega \end{bmatrix}} \cdot \begin{bmatrix} (1-\pi_t)dt \\ \pi_t dt \end{bmatrix}.$$

- For $i \in \{\alpha, \rho\}$,

$$\lambda h_i > s_i > 0.$$

Tasks and Payoffs (cont.)

- Conditional on ω, the expected payoff increment to player i is

$$(1-\pi_t)s_i dt + \pi_t \lambda h_i \omega dt = \boxed{\begin{bmatrix} s_i & \lambda h_i \omega \end{bmatrix}} \cdot \begin{bmatrix} (1-\pi_t)dt \\ \pi_t dt \end{bmatrix}.$$

- For $i \in \{\alpha, \rho\}$,

$$\lambda h_i > s_i > 0.$$

- Preferred allocation coincides if the state is known.

Conflict of Interests

- Let η_i be the (net) **benefit-cost ratio** from the experimentation

$$\eta_i = \frac{\lambda h_i - s_i}{s_i}.$$

Conflict of Interests

- Let η_i be the (net) **benefit-cost ratio** from the experimentation

$$\eta_i = \frac{\lambda h_i - s_i}{s_i}.$$

- Parameters are such that

$$\eta_\alpha > \eta_\rho.$$

Conflict of Interests

- Let η_i be the (net) **benefit-cost ratio** from the experimentation

$$\eta_i = \frac{\lambda h_i - s_i}{s_i}.$$

- Parameters are such that

$$\eta_\alpha > \eta_\rho.$$

- Interpretations:
  - High cost of Principal's resources;
  - Principal's moderate benefit from one out of her many responsibilities;
  - Agent's career advancement as an extra benefit.

Private Information

- Players do not observe the state.
- Agent has private information: his type is his prior belief that the state is 1.
- Agent's type is denoted θ, drawn from $\Theta \equiv [\underline{\theta}, \overline{\theta}] \subset (0, 1)$.
- F is the cdf, f the pdf.

Private Information

- Players do not observe the state.
- Agent has private information: his type is his prior belief that the state is 1.
- Agent's type is denoted θ, drawn from $\Theta \equiv [\underline{\theta}, \overline{\theta}] \subset (0, 1)$.
- F is the cdf, f the pdf.

- Actions and successes are publicly observed.

Policy and Payoffs

- A resource allocation **policy** is a non-anticipative stochastic process $\pi = \{\pi_t\}_{t \geq 0}$.
  - $\pi_t \in [0, 1]$: the fraction allocated to R at time t, which may depend only on the history of events up to t.
- The space of all (mixed) policies is Π.

▸ Formal definition

Examples of Policies

- Allocate all resource to R until a fixed time and switch to S if no success occurs by then;
- Allocate all resource to R until 1st success and then allocate a fixed fraction to R;
- Allocate all resource to R until 2nd success and then switch to S;
- ...

Policy and Payoffs (cont.)

- Players discount payoffs at rate $r > 0$.
- N_t: the number of successes observed up to time t.
- Player i's **payoff** given policy $\pi \in \Pi$ and prior $p_0 \in [0, 1]$ is

$$U_i(\pi, p_0) \equiv \mathbf{E}\left[\int_0^\infty re^{-rt}\left[(1-\pi_t)\, s_i dt + h_i dN_t\right] \middle| \, \pi, p_0\right].$$

Policy and Payoffs (cont.)

- Players discount payoffs at rate $r > 0$.
- N_t: the number of successes observed up to time t.
- Player i's **payoff** given policy $\pi \in \Pi$ and prior $p_0 \in [0,1]$ is

$$U_i(\pi, p_0) \equiv \mathbf{E}\left[\int_0^\infty re^{-rt}\left[(1-\pi_t)\, s_i dt + h_i dN_t\right] \,\middle|\, \pi, p_0\right].$$

- By the Law of Iterated Expectations, $U_i(\pi, p_0)$ can be rewritten as

$$U_i(\pi, p_0) = \mathbf{E}\left[\int_0^\infty re^{-rt}\left[(1-\pi_t)s_i + \pi_t \lambda h_i \omega\right] dt \,\middle|\, \pi, p_0\right].$$

Delegation

- Principal has full commitment and cannot use transfers.
- She determines a delegation contract at time 0.
- By the Revelation Principle, Principal offers a direct mechanism $\boldsymbol{\pi} : \Theta \to \Pi$

$$\sup \int_{\Theta} U_{\rho}(\boldsymbol{\pi}(\theta), \theta) dF(\theta),$$

$$\text{subject to} \quad U_{\alpha}(\boldsymbol{\pi}(\theta), \theta) \geq U_{\alpha}(\boldsymbol{\pi}(\theta'), \theta) \ \ \forall \theta, \theta' \in \Theta.$$

Outline

1. Model
2. Single-player benchmark
3. Characterizing the policy space
4. Main results
5. More general results

Posterior Beliefs

- Given prior p_0 and the history of events up to time t

$$p_t = \mathbf{P}_t[\omega = 1].$$

- Before the first success, p_t satisfies a differential equation

$$\dot{p}_t = -\lambda \pi_t p_t (1 - p_t).$$

- At the first success, p_t jumps to one.

Single Player's Preferred Policy

- Player i's preferred policy is Markov wrt p_t, characterized by a cutoff p_i^* s.t.

$$\pi_t = \begin{cases} 1 & \text{if } p_t > p_i^*, \\ 0 & \text{if } p_t \leq p_i^*. \end{cases}$$

- The cutoff belief is

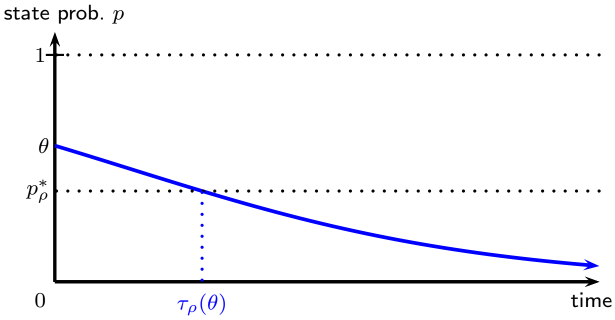
$$p_i^* = \frac{r}{r + (\lambda + r)\eta_i}.$$

- Agent's cutoff is lower than Principal's

$$p_\alpha^* < p_\rho^*.$$

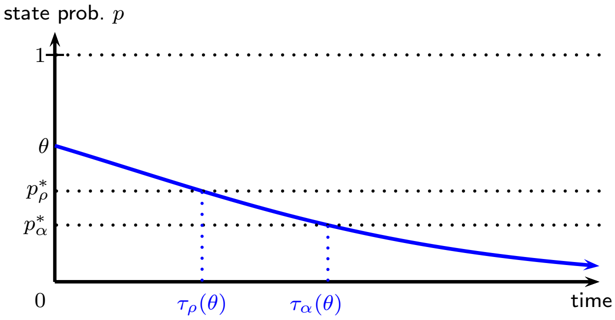
Agency Problem Revisited

- $\tau_i(\theta)$: Player i's preferred stopping time given θ.



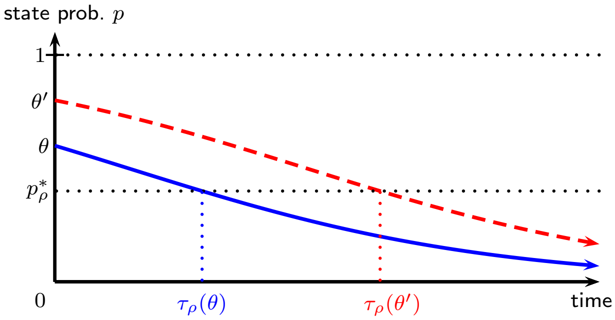
Agency Problem Revisited

- For a given prior, Agent prefers to experiment longer than Principal.



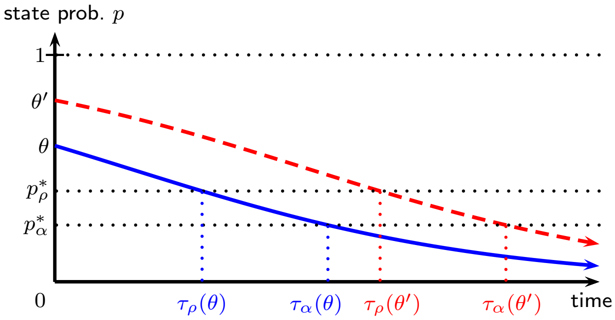
Agency Problem Revisited

- Higher priors warrant longer experimentation.



Agency Problem Revisited

- Lower types (those with lower θ) have incentives to mimic higher types.



Outline

1. Model
2. Single-player benchmark
3. Characterizing the policy space
4. Main results
5. More general results

(Total Expected Discounted) Resource Pair

- For a fixed policy π, define $\mathbf{w}^1(\pi)$ and $\mathbf{w}^0(\pi)$ as follows:

$$\mathbf{w}^1(\pi) \equiv \mathbf{E}\left[\int_0^\infty re^{-rt}\pi_t dt \,\middle|\, \pi, 1\right] \in [0,1]$$
$$\mathbf{w}^0(\pi) \equiv \mathbf{E}\left[\int_0^\infty re^{-rt}\pi_t dt \,\middle|\, \pi, 0\right] \in [0,1].$$

- $\mathbf{w}^1(\pi)$: **(total expected discounted) resource** allocated to R under π in state 1.
- $\mathbf{w}^0(\pi)$: **(total expected discounted) resource** allocated to R under π in state 0.

Summary Statistic for the Payoffs

Lemma 1 (A Policy as a Pair of Numbers)

For a given policy $\pi \in \Pi$ and prior $p_0 \in [0,1]$, player i's payoff can be written as

$$U_i(\pi, p_0) - s_i = \begin{bmatrix} p_0 & 1 - p_0 \end{bmatrix} \cdot \begin{bmatrix} (\lambda h_i - s_i)\, \mathbf{w}^1(\pi) \\ (0 - s_i)\, \mathbf{w}^0(\pi) \end{bmatrix}.$$

▸ Proof

Summary Statistic for the Payoffs

Lemma 1 (A Policy as a Pair of Numbers)

For a given policy $\pi \in \Pi$ and prior $p_0 \in [0,1]$, player i's payoff can be written as

$$U_i(\pi, p_0) - s_i = \begin{bmatrix} p_0 & 1 - p_0 \end{bmatrix} \cdot \begin{bmatrix} (\lambda h_i - s_i)\,\mathbf{w}^1(\pi) \\ (0 - s_i)\,\mathbf{w}^0(\pi) \end{bmatrix}.$$

▸ Proof

$(\mathbf{w}^1(\pi), \mathbf{w}^0(\pi))$ is a summary statistic of π for the payoffs.

Feasible Set

Feasible set Γ: the set of feasible resource pairs

$$\Gamma = \left\{(w^1, w^0) \mid (w^1, w^0) = (\mathbf{w}^1(\pi), \mathbf{w}^0(\pi)),\ \pi \in \Pi\right\}.$$

Characterizing the Feasible Set

$$\hat{w} \in \mathsf{bd}(\Gamma) \iff \exists p \in \mathbf{R}^2, \|p\| = 1, \hat{w} \in \operatorname*{argmax}_{w \in \Gamma} p \cdot w.$$

Characterizing the Feasible Set

$$\hat{w} \in \mathsf{bd}(\Gamma) \iff \exists p \in \mathbf{R}^2, \|p\| = 1, \hat{w} \in \operatorname*{argmax}_{w \in \Gamma} p \cdot w.$$

Characterizing the Feasible Set

$$\hat{w} \in \text{bd}(\Gamma) \iff \exists p \in \mathbf{R}^2, \|p\| = 1, \hat{w} \in \operatorname*{argmax}_{w \in \Gamma} p \cdot w.$$

Characterizing the Feasible Set

$$\hat{w} \in \mathsf{bd}(\Gamma) \iff \exists p \in \mathbf{R}^2, \|p\| = 1, \hat{w} \in \operatorname*{argmax}_{w \in \Gamma} p \cdot w.$$

Characterizing the Feasible Set

$$\hat{w} \in \mathsf{bd}(\Gamma) \iff \exists p \in \mathbf{R}^2, \|p\| = 1, \hat{w} \in \operatorname*{argmax}_{w \in \Gamma} p \cdot w.$$

Characterizing the Feasible Set

$$\hat{w} \in \mathsf{bd}(\Gamma) \iff \exists p \in \mathbf{R}^2, \|p\| = 1, \hat{w} \in \underset{w \in \Gamma}{\operatorname{argmax}}\, p \cdot w.$$

Lemma 2 (Feasible Set)

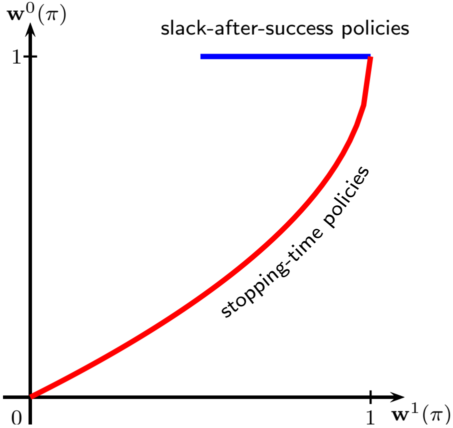
$\Gamma = \mathsf{co}\left\{(\mathbf{w}^1(\pi), \mathbf{w}^0(\pi)), \pi \in \Pi^M\right\}$, where Π^M are Markov policies (wrt p).

Proof

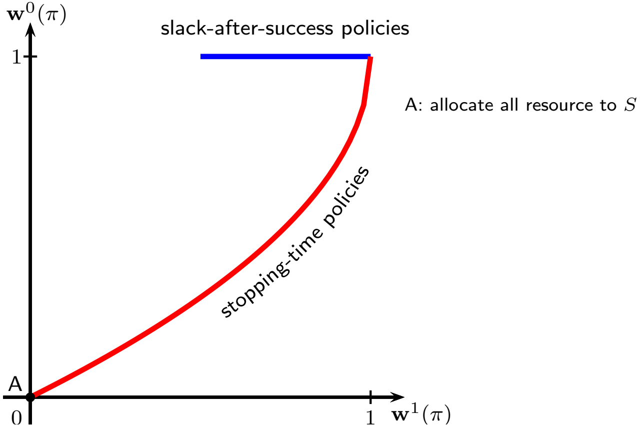
Canonical Markov Policies: Poisson Conclusive News

- Stopping-time policies (lower-cutoff Markov policies)
  - allocate all resource to R until a fixed time;
  - if at least one success occurs by then, allocate all resource to R forever;
  - otherwise, switch to S.
- Slack-after-success policies (upper-cutoff Markov policies)
  - allocate all resource to R until the first success;
  - then allocate a fixed fraction to R.

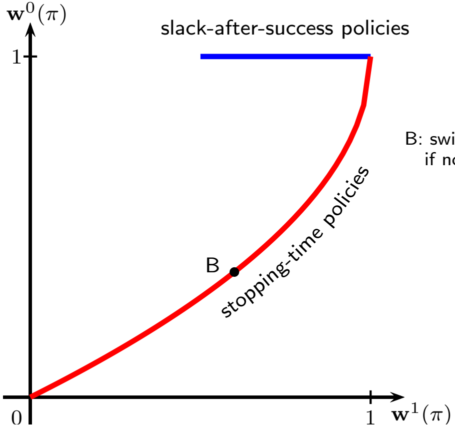
Canonical Markov Policies: Poisson Conclusive News



Canonical Markov Policies: Poisson Conclusive News

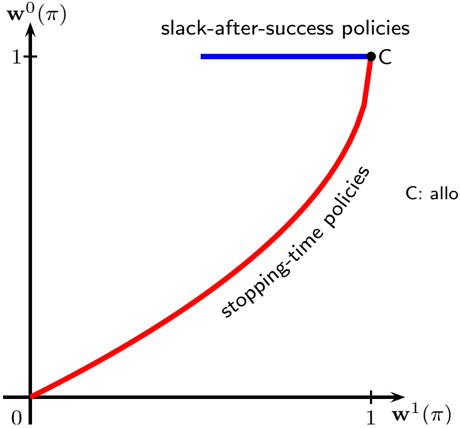


Canonical Markov Policies: Poisson Conclusive News



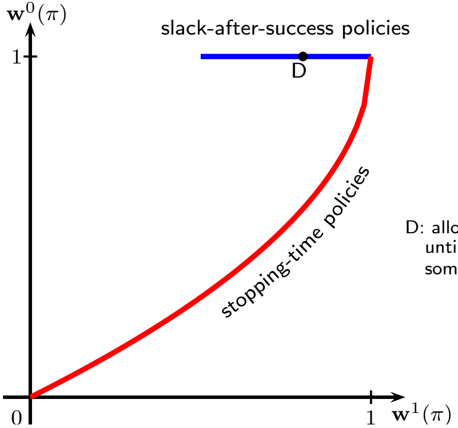
B: switch to S at some fixed time
if no success occurs

Canonical Markov Policies: Poisson Conclusive News



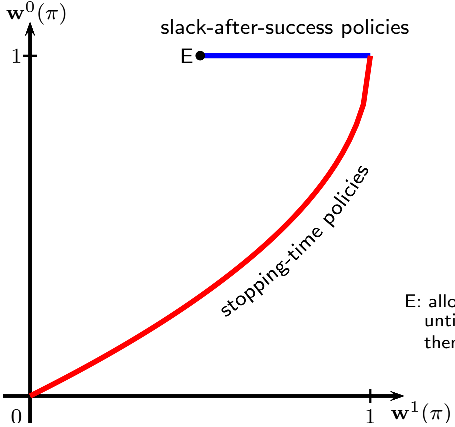
C: allocate all resource to R

Canonical Markov Policies: Poisson Conclusive News



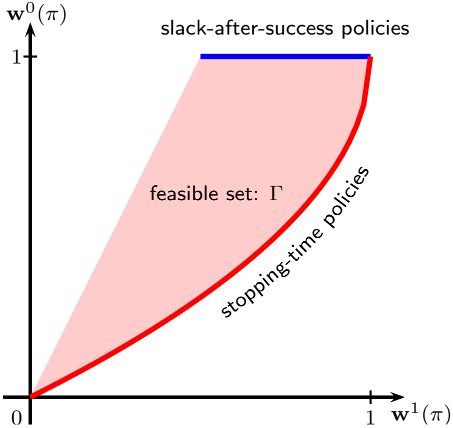
D: allocate all resource to R until 1st success; then allocate some fixed fraction to R

Canonical Markov Policies: Poisson Conclusive News



E: allocate all resource to R
until 1st success;
then switch to S

Feasible Set: Poisson Conclusive News



Feasible Set: Poisson Conclusive News

Lemma 3 (Feasible Set: Poisson Conclusive News)

The feasible set is the convex hull of the image of stopping-time and slack-after-success policies.

Preferences over Feasible Pairs

- Player i's payoff given π and θ is

$$U_i(\pi, \theta) - s_i = \left[\theta \eta_i \mathbf{w}^1(\pi) - (1-\theta)\mathbf{w}^0(\pi)\right] \cdot s_i.$$

Preferences over Feasible Pairs

- Player i's payoff given π and θ is

$$U_i(\pi,\theta) - s_i = \boxed{\left[\theta\eta_i \mathbf{w}^1(\pi) - (1-\theta)\mathbf{w}^0(\pi)\right]} \cdot s_i.$$

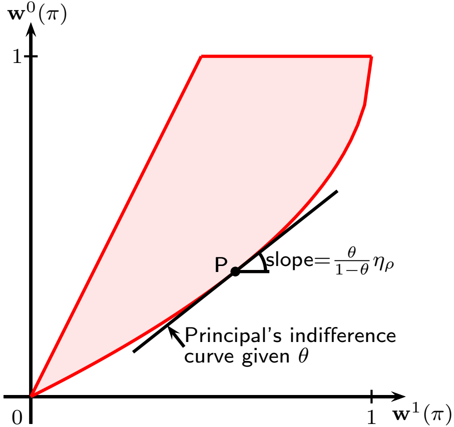
Preferences over Feasible Pairs

- Player i's payoff given π and θ is

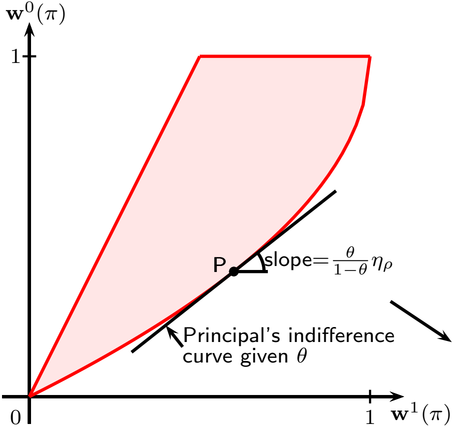
$$U_i(\pi, \theta) - s_i = \boxed{\left[\theta \eta_i \mathbf{w}^1(\pi) - (1-\theta)\mathbf{w}^0(\pi)\right]} \cdot s_i.$$

- Player i's preferences over $(w^1, w^0) \in \Gamma$ are determined by
  - θ: the prior belief that the state is 1;
  - η_i: the benefit-cost ratio from the experimentation.

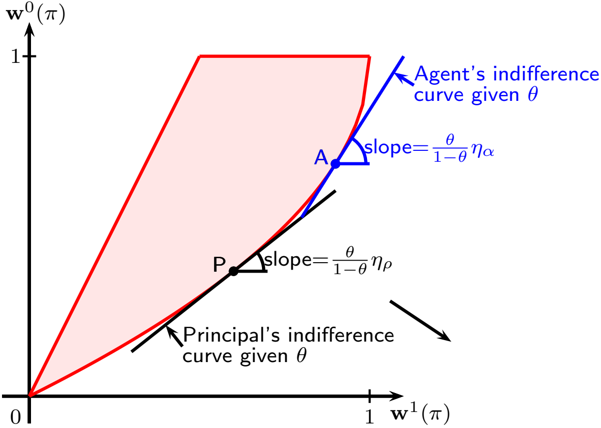
Preferences over Feasible Pairs: Indifference Curves



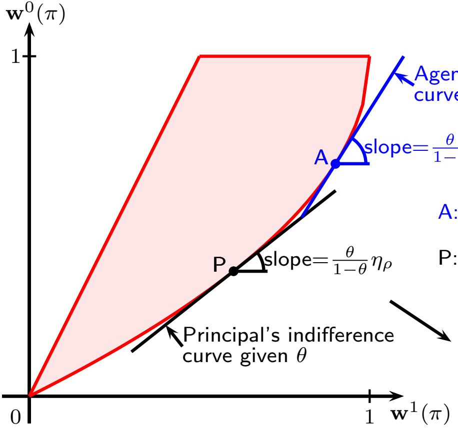
Preferences over Feasible Pairs: Indifference Curves



Preferences over Feasible Pairs: Indifference Curves



Preferences over Feasible Pairs: Indifference Curves



A: $(w^1_\alpha(\theta), w^0_\alpha(\theta))$

P: $(w^1_\rho(\theta), w^0_\rho(\theta))$

Delegation Problem Reformulated

- Replace policy space Π with feasible set Γ:

$$(w^1, w^0) \in \Gamma \Longrightarrow \theta\eta_i w^1 - (1-\theta)w^0.$$

Delegation Problem Reformulated

- Replace policy space Π with feasible set Γ:

$$(w^1, w^0) \in \Gamma \Longrightarrow \theta\eta_i w^1 - (1-\theta)w^0.$$

- Principal offers a direct mechanism $(w^1, w^0) : \Theta \to \Gamma$

$$\max \int_\Theta \left(\theta\eta_\rho w^1(\theta) - (1-\theta)w^0(\theta)\right) dF(\theta),$$

$$\text{subject to} \quad \frac{\theta\eta_\alpha}{1-\theta}w^1(\theta) - w^0(\theta) \geq \frac{\theta\eta_\alpha}{1-\theta}w^1(\theta') - w^0(\theta') \quad \forall\theta, \theta' \in \Theta.$$

Delegation Problem Reformulated

- Replace policy space Π with feasible set Γ:

$$(w^1, w^0) \in \Gamma \Longrightarrow \theta \eta_i w^1 - (1-\theta) w^0.$$

- Principal offers a direct mechanism $(w^1, w^0) : \Theta \to \Gamma$

$$\max \int_\Theta \left(\theta \eta_\rho w^1(\theta) - (1-\theta) w^0(\theta)\right) dF(\theta),$$

$$\text{subject to} \quad \frac{\theta \eta_\alpha}{1-\theta} w^1(\theta) - w^0(\theta) \geq \frac{\theta \eta_\alpha}{1-\theta} w^1(\theta') - w^0(\theta') \ \ \forall \theta, \theta' \in \Theta.$$

- Payoff parameters $\eta_\alpha > \eta_\rho$; feasible set Γ ; type distribution F.

Delegation Problem Reformulated

- Replace policy space Π with feasible set Γ:

$$(w^1, w^0) \in \Gamma \Longrightarrow \theta\eta_i w^1 - (1-\theta)w^0.$$

- Principal offers a direct mechanism $(w^1, w^0) : \Theta \to \Gamma$

$$\max \int_\Theta \left(\theta\eta_\rho w^1(\theta) - (1-\theta)w^0(\theta)\right) dF(\theta),$$

$$\text{subject to} \quad \frac{\theta\eta_\alpha}{1-\theta} w^1(\theta) - w^0(\theta) \geq \frac{\theta\eta_\alpha}{1-\theta} w^1(\theta') - w^0(\theta') \;\; \forall \theta, \theta' \in \Theta.$$

- Payoff parameters $\eta_\alpha > \eta_\rho$; feasible set Γ; type distribution F.

Outline

1. Model
2. Single-player benchmark
3. Characterizing the policy space
4. Main results
5. More general results

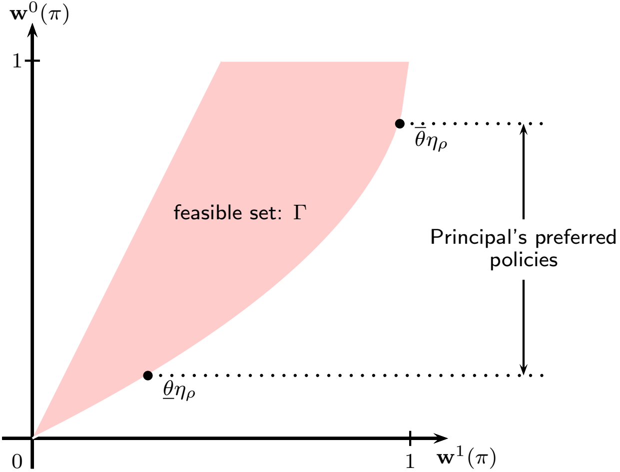
The Cutoff Rule

Definition 1

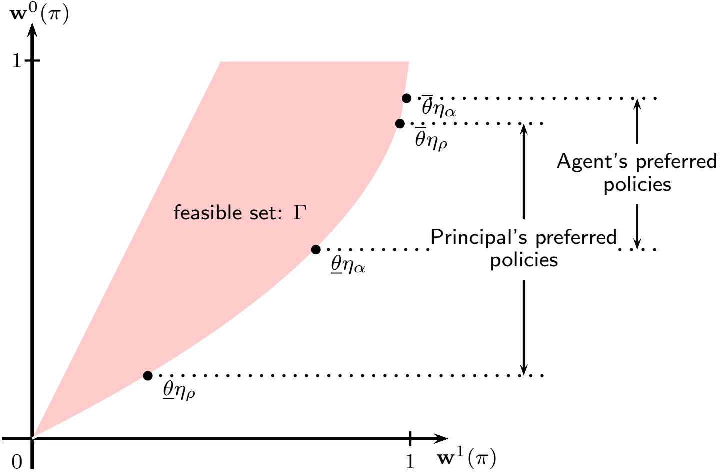
The **cutoff rule** is the contract (w^1, w^0) s.t.

$$(w^1(\theta), w^0(\theta)) = \begin{cases} (w^1_\alpha(\theta), w^0_\alpha(\theta)) & \text{if } \theta \leq \theta^*, \\ (w^1_\alpha(\theta^*), w^0_\alpha(\theta^*)) & \text{if } \theta > \theta^*. \end{cases}$$

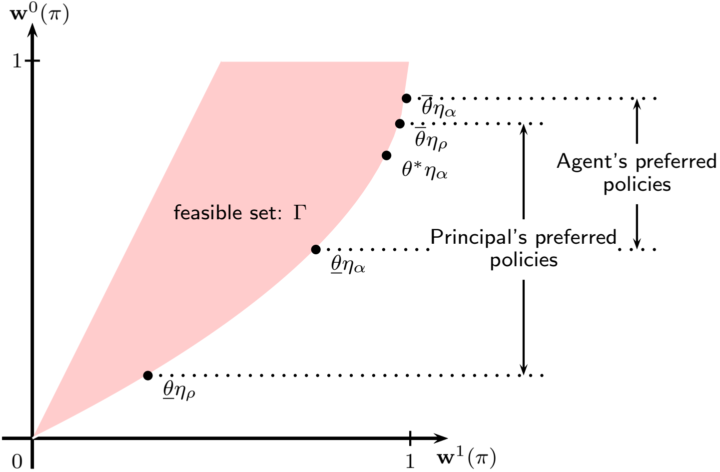
Delegation Set under Cutoff Rule



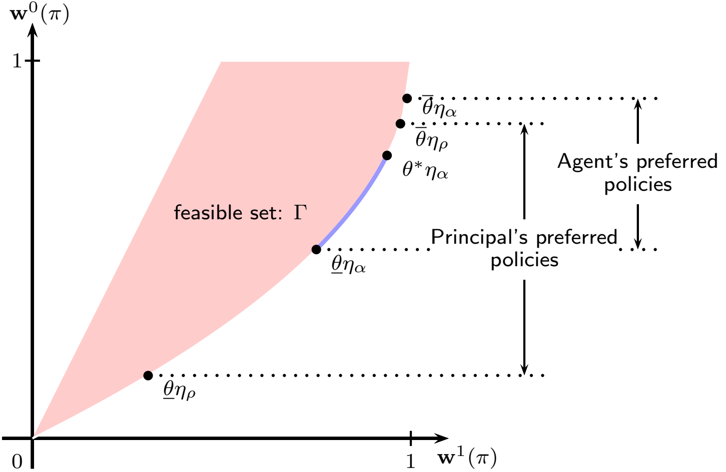
Delegation Set under Cutoff Rule



Delegation Set under Cutoff Rule



Delegation Set under Cutoff Rule



Optimality

Main assumption

For all $\theta \leq \theta^*$, the following condition is satisfied:

$$\frac{\eta_\alpha}{\eta_\alpha - \eta_\rho} \geq (3\theta - 1) - \frac{f'(\theta)}{f(\theta)}\theta(1-\theta).$$

Proposition 1

The cutoff rule is optimal if the main assumption holds.

Implementing the Cutoff Rule

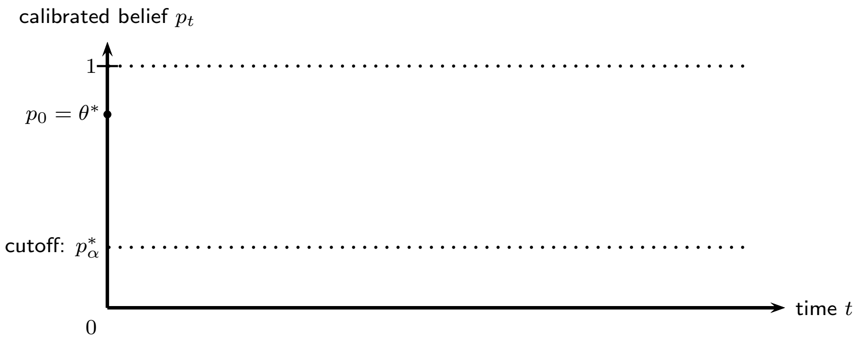
Calibrated belief (p_t)

- prior belief $p_0 = \theta^*$;
- without any success, it drifts down according to $\dot{p}_t = -\lambda \pi_t p_t (1 - p_t)$;
- upon the first success, it jumps to one.

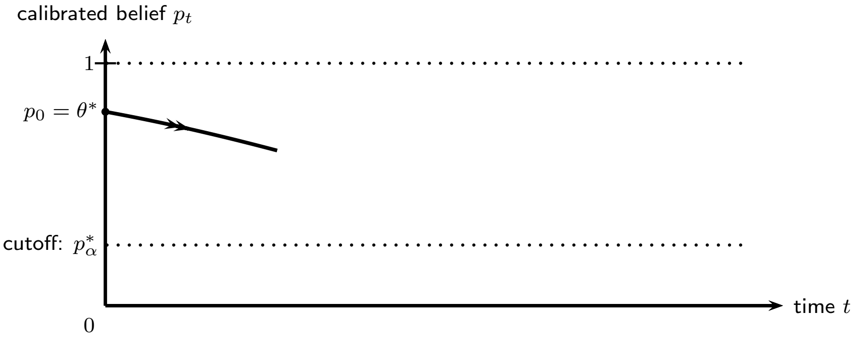
Behavior

- a cutoff imposed at p^*_α;
- Agent has full flexibility if the belief stays above the cutoff;
- Agent is required to stop once the cutoff is reached.

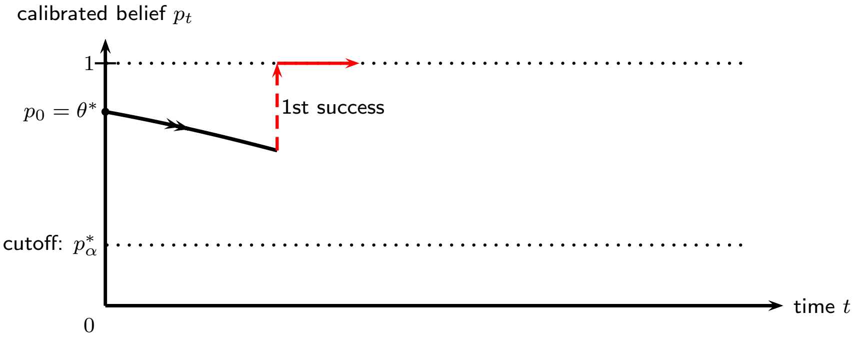
Implementing the Cutoff Rule (cont.)



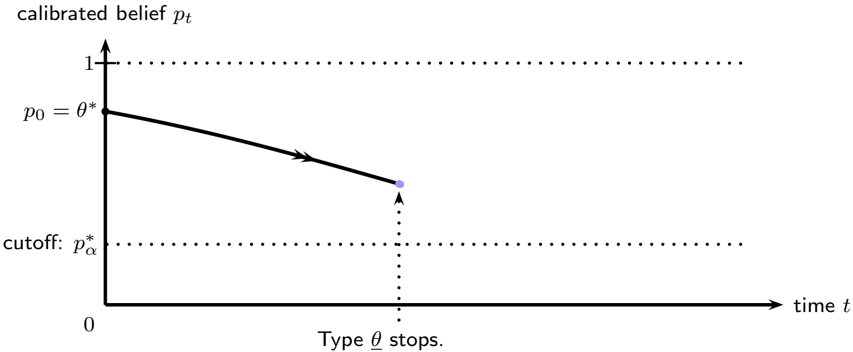
Implementing the Cutoff Rule (cont.)



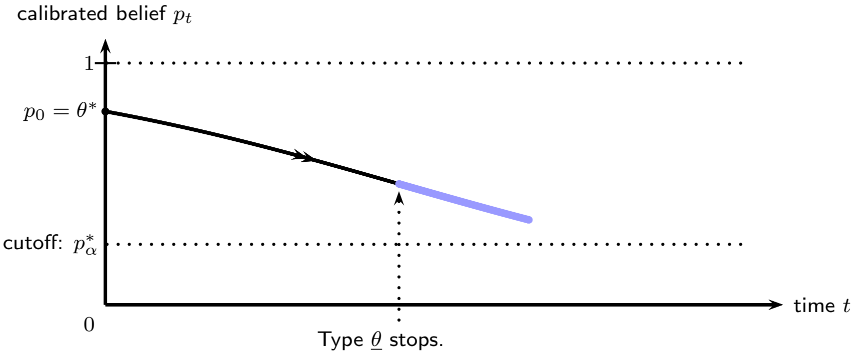
Implementing the Cutoff Rule (cont.)



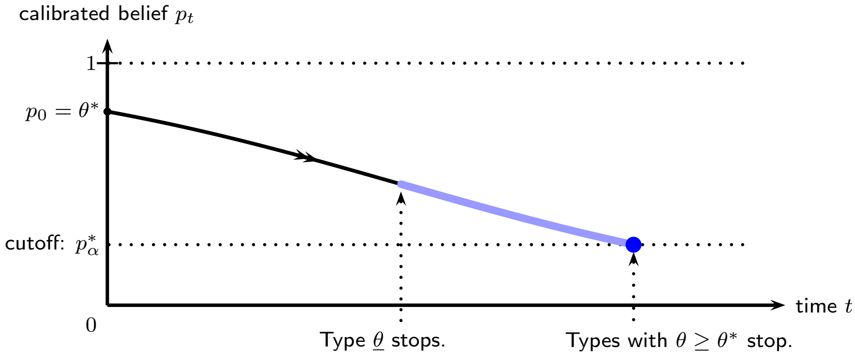
Implementing the Cutoff Rule (cont.)



Implementing the Cutoff Rule (cont.)



Implementing the Cutoff Rule (cont.)



Time Consistency

Definition 2

Fix a (direct or indirect) mechanism. It is **time-consistent** if Principal finds it optimal to fulfill the mechanism after any history on path.

▸ Formal definition

Time Consistency

Definition 2

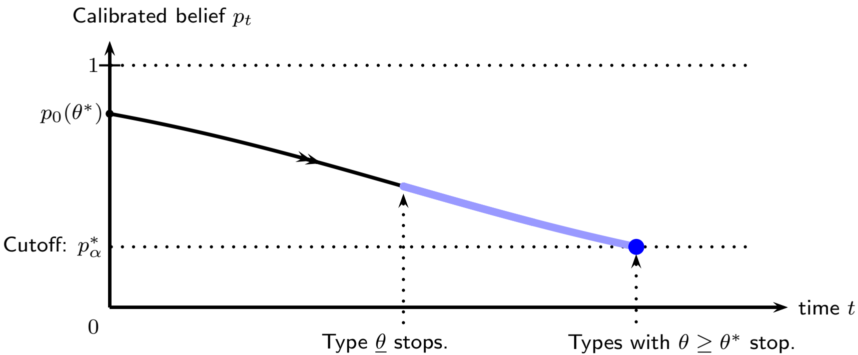
Fix a (direct or indirect) mechanism. It is **time-consistent** if Principal finds it optimal to fulfill the mechanism after any history on path.

Formal definition

Proposition 2

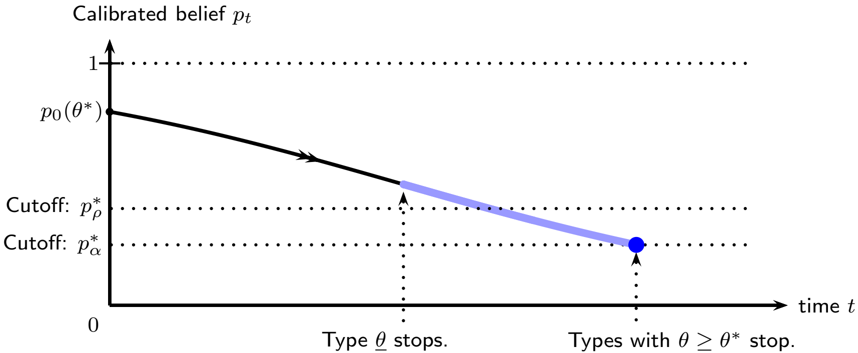
The cutoff rule is time-consistent if the main assumption holds.

Time Consistency: Principal's Posterior Belief



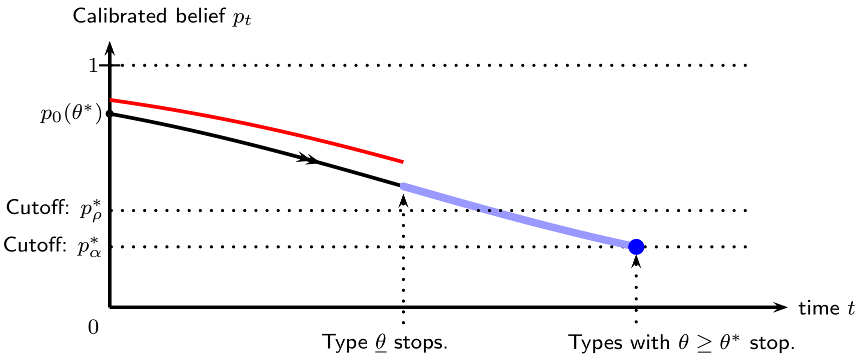
More general results

Time Consistency: Principal's Posterior Belief



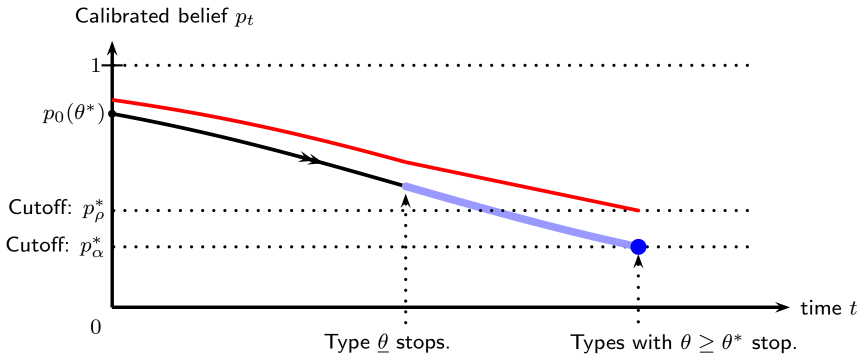
More general results

Time Consistency: Principal's Posterior Belief



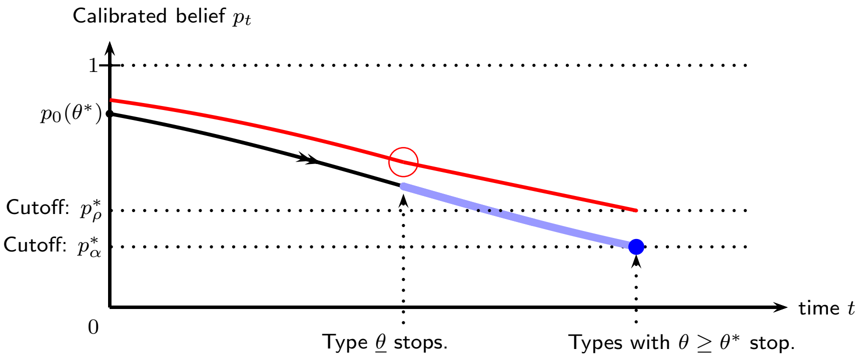
▸ More general results

Time Consistency: Principal's Posterior Belief



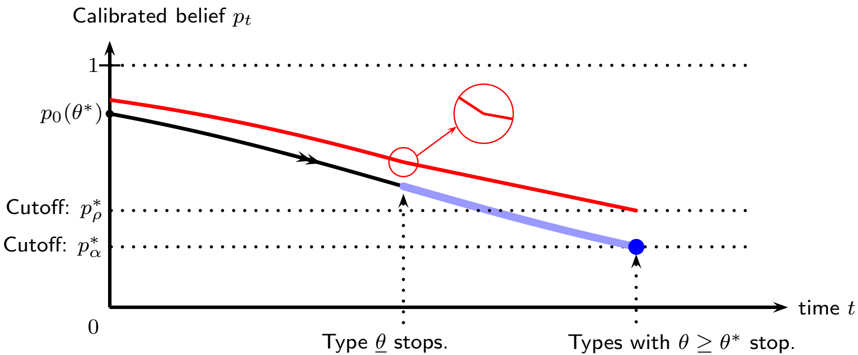
▸ More general results

Time Consistency: Principal's Posterior Belief



More general results

Time Consistency: Principal's Posterior Belief



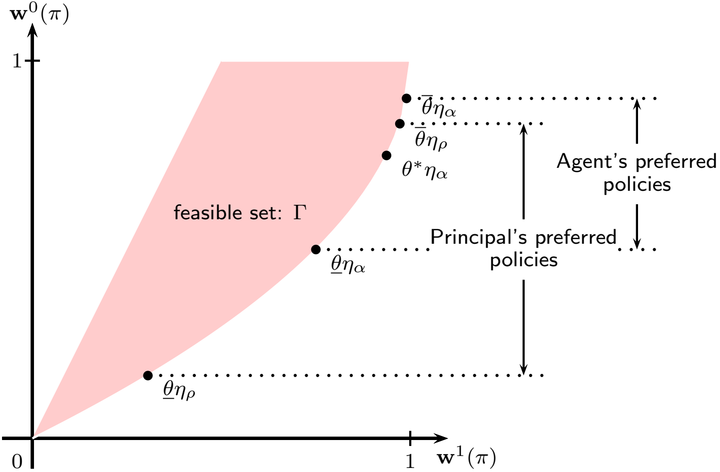
▸ More general results

The Cutoff Type

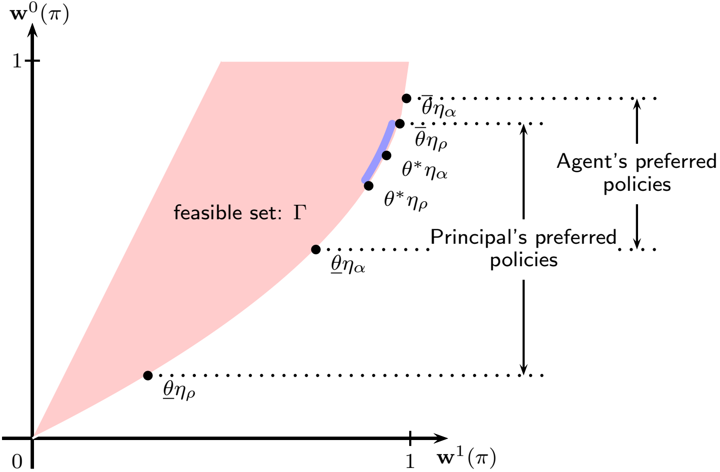
The **cutoff type** θ^*: the lowest value in Θ s.t.

- Agent's preferred policy given θ^* equals Principal's preferred policy if she believes that $\theta \geq \theta^*$.
- For any $\hat{\theta} > \theta^*$, Agent's preferred policy given $\hat{\theta}$ is above Principal's preferred policy if she believes that $\theta \geq \hat{\theta}$.

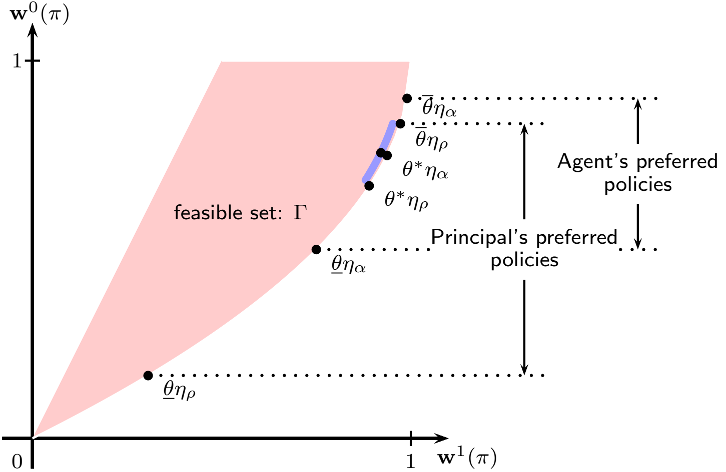
The Cutoff Type (cont.)



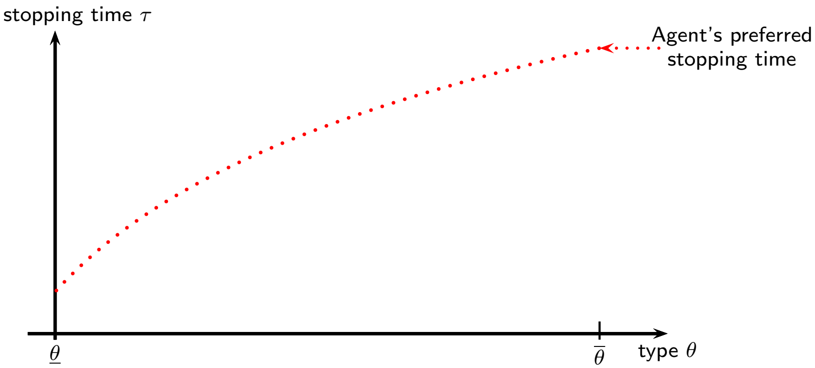
The Cutoff Type (cont.)



The Cutoff Type (cont.)

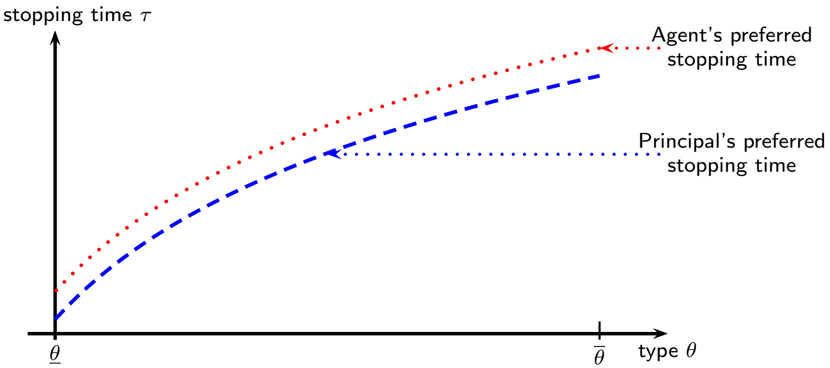


Over- and Under-Experimentation



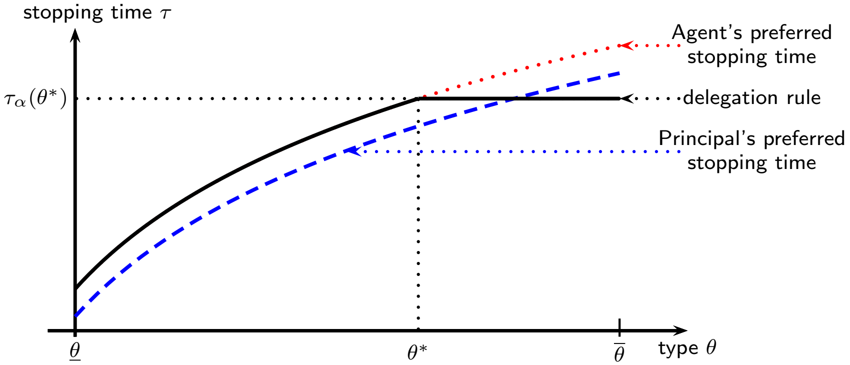
More general results

Over- and Under-Experimentation



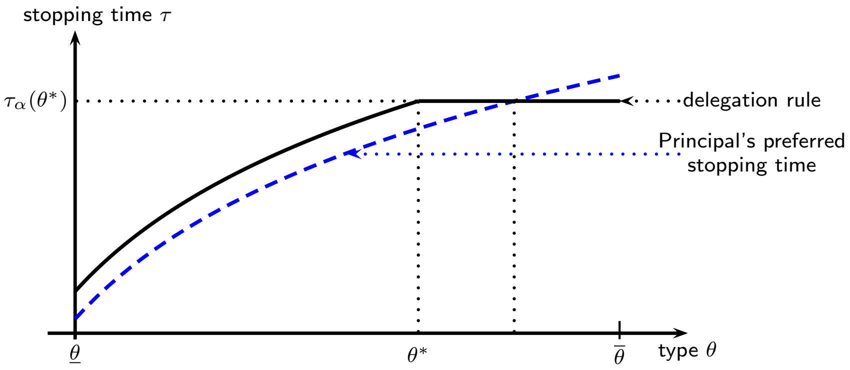
More general results

Over- and Under-Experimentation



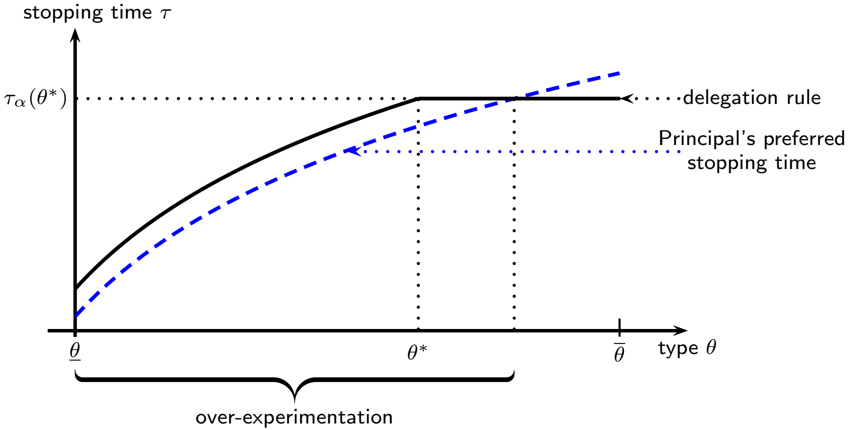
▸ More general results

Over- and Under-Experimentation



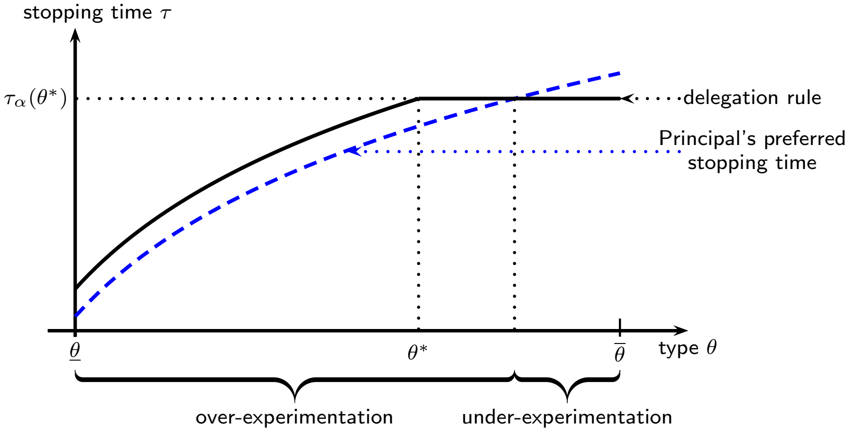
▸ More general results

Over- and Under-Experimentation



▸ More general results

Over- and Under-Experimentation

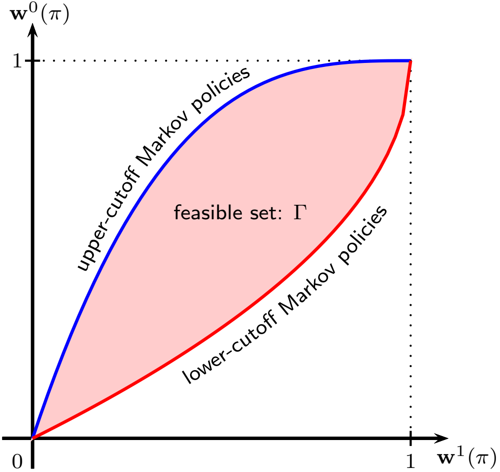


More general results

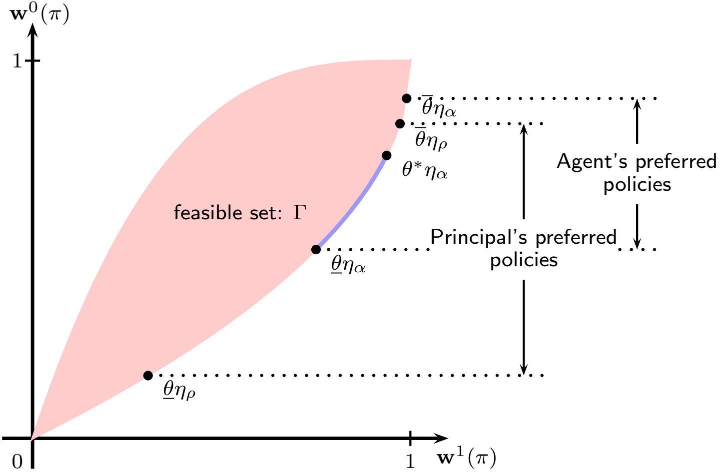
Outline

1. Model
2. Single-player benchmark
3. Characterizing the policy space
4. Main results
5. More general results

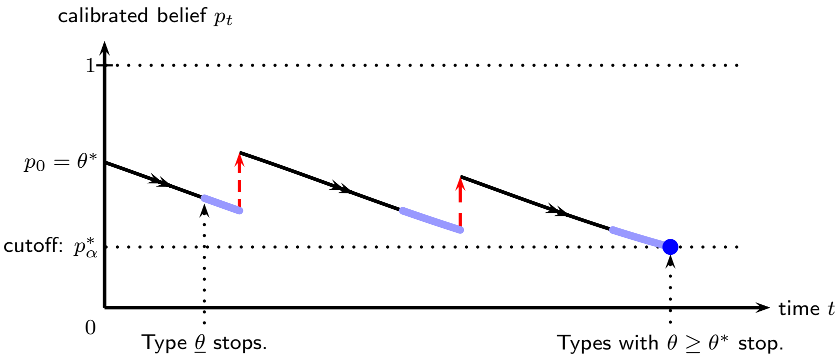
Feasible Set: Poisson Inconclusive News



The Cutoff Rule: Poisson Inconclusive News



Implementation: Poisson Inconclusive News



Sliding Deadline: Poisson Inconclusive News

- Principal initially sets a deadline for experimentation
- Whenever a success realizes, the deadline is extended.
- Agent is free to switch to S before the deadline.
- When the deadline is reached, Agent is required to switch to S.

The end

Lévy Bandits

Proposition 3

The cutoff rule is optimal if the main assumption holds.

Proposition 4

The cutoff rule is time-consistent if the main assumption holds.

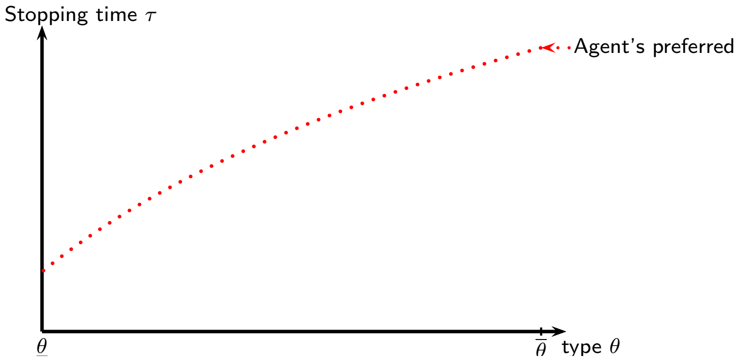
▸ The end

Optimal Contract with Transfers

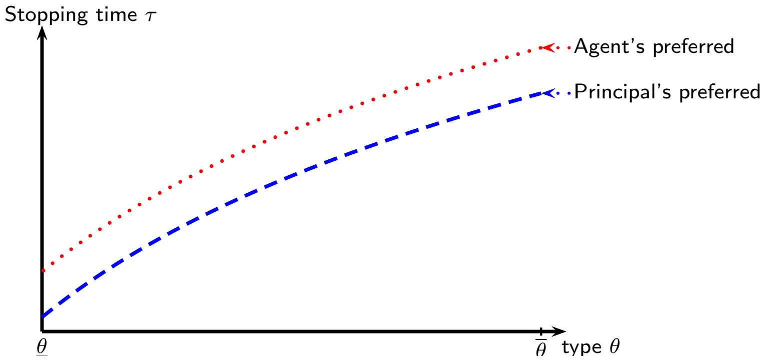
- The principal can make transfers to the agent.
- The agent is protected by limited liability.
- For each type, the principal specifies an experimentation policy and a transfer scheme,

$$\left(w^1, w^0; t^1, t^0\right).$$

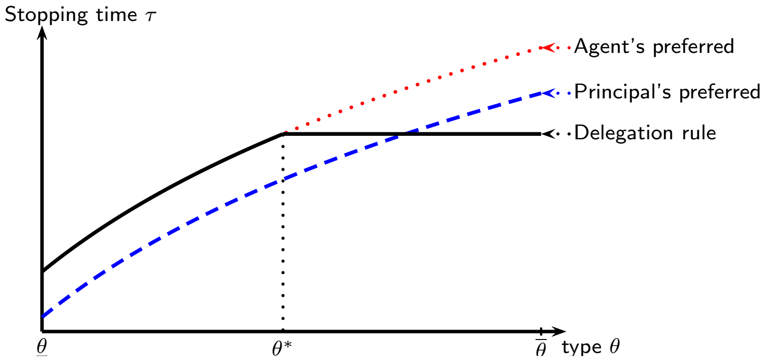
Optimal Contract with Transfers



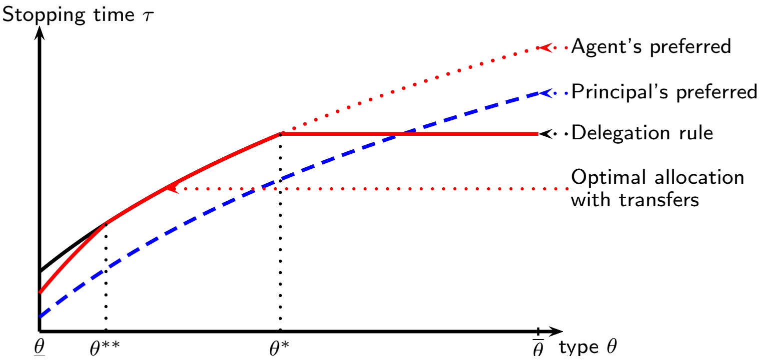
Optimal Contract with Transfers



Optimal Contract with Transfers



Optimal Contract with Transfers



Implications and Applications

- A (sliding) deadline should be in place as a safeguard against abuse of resources. The continuation of a project is permitted only upon demonstrated successes.
- Agent should have full flexibility over resource allocation before the (sliding) deadline is reached.

Implications and Applications

- A (sliding) deadline should be in place as a safeguard against abuse of resources. The continuation of a project is permitted only upon demonstrated successes.
- Agent should have full flexibility over resource allocation before the (sliding) deadline is reached.

- Google: the once highly publicized and well-funded Google Wave was canceled in August 2010 as it failed to achieve the goal set by Google executives before then.

Future Work

- Asymmetric learning.
- Multi-dimensional hidden information.
- Allocation vs. investment.

Thank you!

Appendix

A Policy is a Non-Anticipative Process

- Suppose the process L is a Lévy process L^1 with probability $p \in (0,1)$ and L^0 with probability $1-p$.
- Let $\mathcal{F}_t^L$ be the sigma-algebra generated by the process $(L(s))_{s \le t}$.
- Then it is required that the process π satisfies that $\{\int_0^t \pi_s ds \le t'\} \in \mathcal{F}_{t'}^L$, for any $t, t' \in [0, \infty)$.

Back

Mixed Policies

- I define mixed policies following Aumann (1964).
- Let Π^* be the set of all pure policies.
- I define mixed policies as measurable functions $\hat{\pi} : [0,1] \to \Pi^*$.
- According to $\hat{\pi}$, a value $x \in [0,1]$ is drawn uniformly from $[0,1]$ and then the pure policy $\hat{\pi}(x)$ is implemented.
- Stochastic mechanisms are measurable functions $\hat{\boldsymbol{\pi}} : \Theta \times [0,1] \to \Pi^*$.

Back

A Policy as a Pair of Numbers

Player i's payoff given policy $\pi \in \Pi$ and prior $p_0 \in [0,1]$ is

$$\begin{aligned} U_i(\pi, p_0) &= \mathbf{E}\left[\int_0^\infty re^{-rt}\left[(1-\pi_t)s_i + \pi_t\lambda^\omega h_i\right]dt \,\middle|\, \pi, p_0\right] \\ &= p_0\mathbf{E}\left[\int_0^\infty re^{-rt}\left[s_i + \pi_t\left(\lambda^1 h_i - s_i\right)\right]dt \,\middle|\, \pi, 1\right] \\ &+ (1-p_0)\mathbf{E}\left[\int_0^\infty re^{-rt}\left[s_i + \pi_t\left(\lambda^0 h_i - s_i\right)\right]dt \,\middle|\, \pi, 0\right] \\ &= p_0\left(\lambda^1 h_i - s_i\right)\mathbf{E}\left[\int_0^\infty re^{-rt}\pi_t dt \,\middle|\, \pi, 1\right] \\ &+ (1-p_0)\left(\lambda^0 h_i - s_i\right)\mathbf{E}\left[\int_0^\infty re^{-rt}\pi_t dt \,\middle|\, \pi, 0\right] + s_i \\ &= p_0\left(\lambda^1 h_i - s_i\right)\mathbf{w}^1(\pi) + (1-p)\left(\lambda^0 h_i - s_i\right)\mathbf{w}^0(\pi) + s_i. \end{aligned}$$

Back

Characterization of Feasible Set

- Given $\gamma = (\gamma^1, \gamma^0) \in \mathbf{R}^2$, define the supremum score in direction γ and the associated half space as

$$K(\gamma) \equiv \sup_{\pi \in \Pi} \left[\gamma^1 \mathbf{w}^1(\pi) + \gamma^0 \mathbf{w}^0(\pi)\right],$$
$$\mathcal{H}(\gamma) \equiv \left\{v \in \mathbf{R}^2 : \gamma \cdot v \leq K(\gamma)\right\}.$$

- Define the intersection of all half spaces as $\mathcal{H} \equiv \cap_{\gamma \in \mathbf{R}^2} \mathcal{H}(\gamma)$.
- Since $\Gamma \subset \mathcal{H}(\gamma)$ for any γ, it follows that $\Gamma \subset \mathcal{H}$.
- The feasible set Γ is convex given that the policy space Π and hence Γ are convexified.
- It follows that $\Gamma = \mathcal{H}$.
- Since the extreme points of $\mathcal{H}$ are given by Markov policies, this completes the proof.

Back

Formal Definition of Time Consistency

- A history of length t on path is $h^t = ((\pi_s)_{s \le t}, (N_s)_{s \le t})$.
- The set of histories of length t on path is denoted H^t. The set of all histories on path is $H = \cup_{t \ge 0} H^t$.
- Let $F(h^t)$ be the cdf of the agent's belief of state 1 at time t after history h^t with support $\Theta(h^t)$.
- A delegation rule $C : \Theta \to \Pi$ admits a time-consistent implementation if for any h^t on path

$$C(h^t) \in \underset{\boldsymbol{\pi}(\cdot)}{\operatorname{argmax}} \int_{\Theta(h^t)} U_\rho(\theta, \boldsymbol{\pi}(\theta)) dF(h^t),$$

$$\text{subject to} \quad U_\alpha(\theta, \boldsymbol{\pi}(\theta)) \ge U_\alpha(\theta, \boldsymbol{\pi}(\theta')) \ \ \forall \theta, \theta' \in \Theta(h^t).$$

Back