Learning, Experimentation, Information Design

Johannes Hörner[1], Andy Skrzypacz[2]

[1]Yale

[2]Stanford GSB

World Congress, Montréal

We will survey work over the last decade which involves

1. Dynamics.
2. Strategic Interaction.
3. Learning.

We will survey work over the last decade which involves

1. Dynamics.
2. Strategic Interaction.
3. Learning.

Leave out work that involves transfers (principal-agent models, dynamic mechanism design, prizes/wages in experimentation,...)

We will survey work over the last decade which involves

1. Dynamics.
2. Strategic Interaction.
3. Learning.

Leave out work that involves transfers (principal-agent models, dynamic mechanism design, prizes/wages in experimentation,...)

In this talk, will focus on a few themes:

1. Strategic experimentation.
2. Information design.
3. Delegation.

We will survey work over the last decade which involves

1. Dynamics.
2. Strategic Interaction.
3. Learning.

Leave out work that involves transfers (principal-agent models, dynamic mechanism design, prizes/wages in experimentation,. . .)

In this talk, will focus on a few themes:

1. Strategic experimentation.
2. Information design.
3. Delegation.

I relegate to the survey related themes.

Strategic Bandits

Bolton&Harris (1999)
Keller, Rady and Cripps (2005)
Keller&Rady (2015)

Consider a continuous-time game with:

- Players $i = 1, \dots, n$.
- Actions $k_t^i \in [0, 1]$: fraction allocated to risky (R) arm.
- Common binary state of the world $\omega \in \{G, B\}$.

Consider a continuous-time game with:

Players $i = 1, \ldots, n$.
Actions $k_t^i \in [0, 1]$: fraction allocated to risky (R) arm.
Common binary state of the world $\omega \in \{G, B\}$.

Realized payoff:

$$\int_0^\infty re^{-rt}(h\mathrm{d}N_{G,t}^i - \ell\mathrm{d}N_{B,t}^i),$$

where $N_{G,t}^i, N_{B,t}^i$ are Poisson processes with intensity

$$k_t^i \lambda_G \mathbf{1}_{\{\omega=G\}}, \quad k_t^i \lambda_B \mathbf{1}_{\{\omega=B\}},$$

conditionally independent across players. **Conclusive** news.

Pure information externalities. Observable actions, outcomes.

This is a bandit model.

- If $\lambda_G > \lambda_B$: no news is bad, belief $p_t = \mathbf{P}[\omega = G]$ goes down (**Good News**).
- If $\lambda_G < \lambda_B$: no news is good, p_t goes up (**Bad News**).

This is a bandit model.

- If $\lambda_G > \lambda_B$: no news is bad, belief $p_t = \mathbf{P}[\omega = G]$ goes down (**Good News**).
- If $\lambda_G < \lambda_B$: no news is good, p_t goes up (**Bad News**).

Accordingly, **first-best** is a threshold policy s.t.:

- Good News: Play Risky $(k_t^i = 1 \forall i, t)$ iff $p_t \geq \underline{p}$.
- Bad News: Play Risky iff $p_t \geq \overline{p}$.

This is a bandit model.

If $\lambda_G > \lambda_B$: no news is bad, belief $p_t = \mathbf{P}[\omega = G]$ goes down (**Good News**).

If $\lambda_G < \lambda_B$: no news is good, p_t goes up (**Bad News**).

Accordingly, **first-best** is a threshold policy s.t.:

Good News: Play Risky ($k_t^i = 1 \forall i, t$) iff $p_t \geq \underline{p}$.

Bad News: Play Risky iff $p_t \geq \overline{p}$.

On-path behaviors differ markedly:

Good News: experimentation stops **unless** good news occurs.

Bad News: experimentation goes on **unless** bad news occurs.

But this is a **strategic** problem.

But this is a **strategic** problem.

Equilibrium **payoff** uninteresting:

Smaller than in first-best, by definition.
Larger than by oneself, because information can be ignored.

But this is a **strategic** problem.

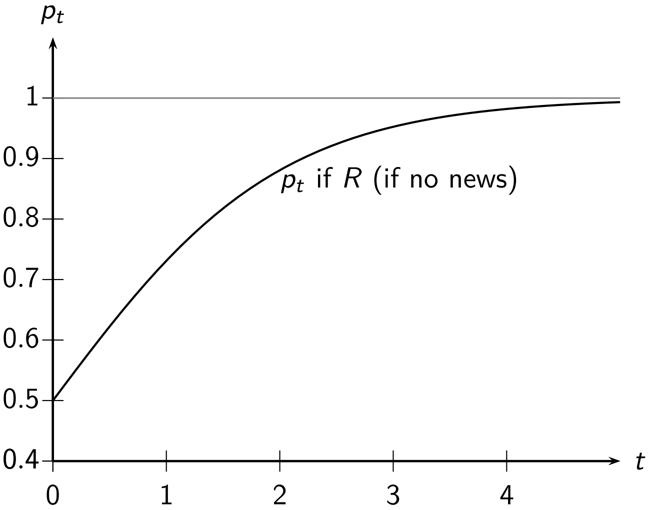
Equilibrium **payoff** uninteresting:

Smaller than in first-best, by definition.
Larger than by oneself, because information can be ignored.

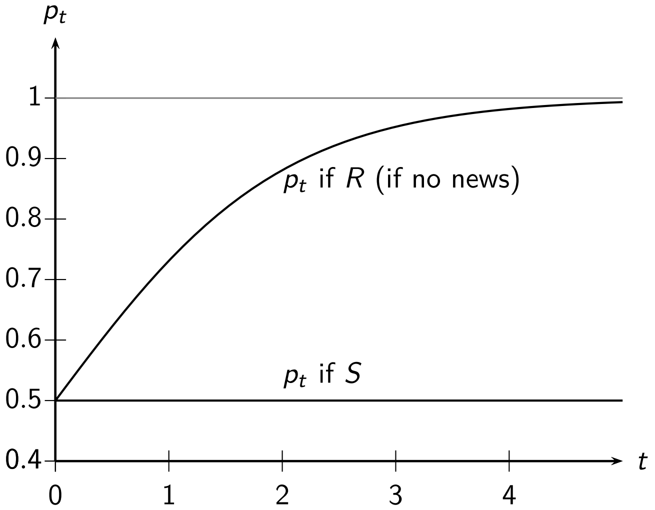
Experimentation **amount** more interesting:

Does each player experiment more or less than on his own?

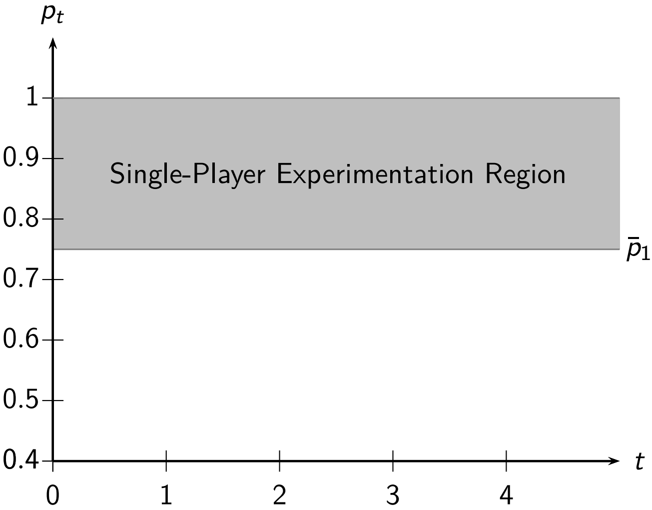
Bad News: Best-Reply



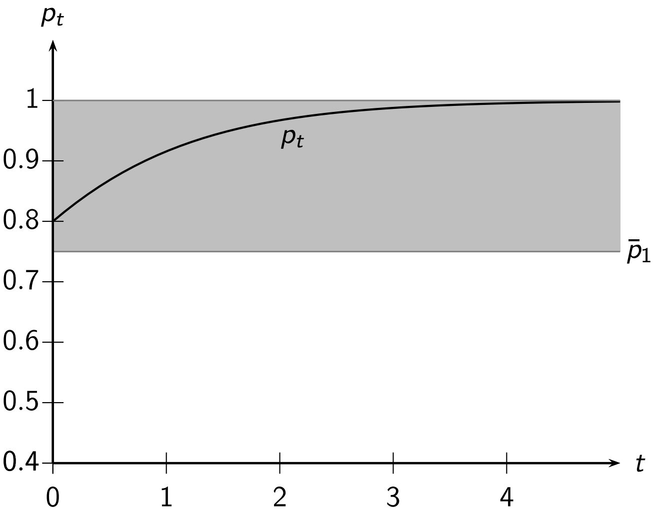
Bad News: Best-Reply



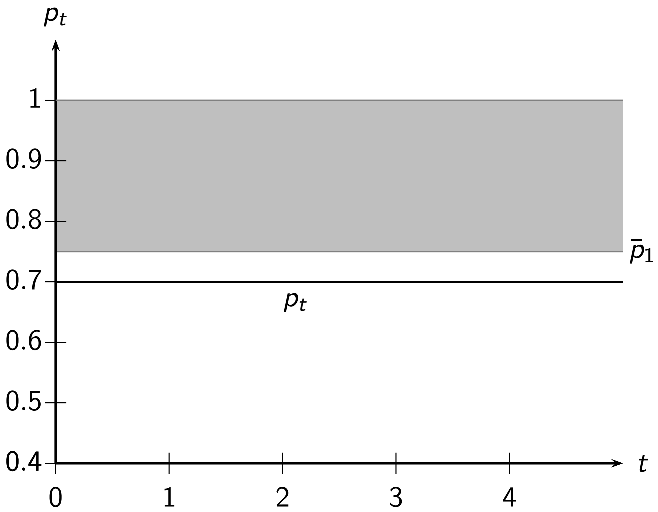
Bad News: Best-Reply



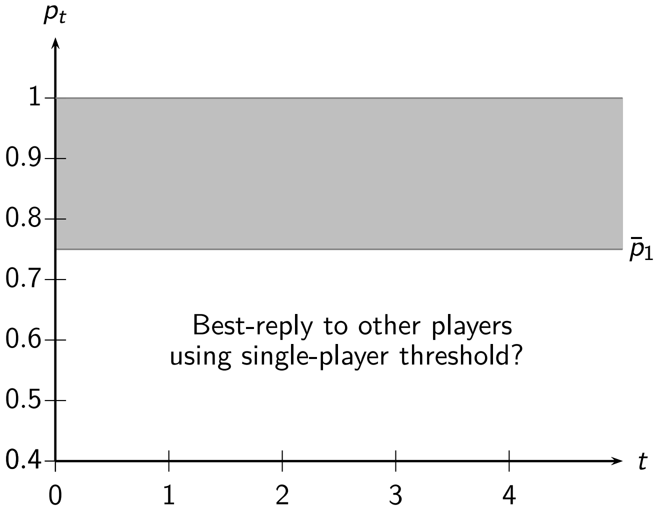
Bad News: Best-Reply



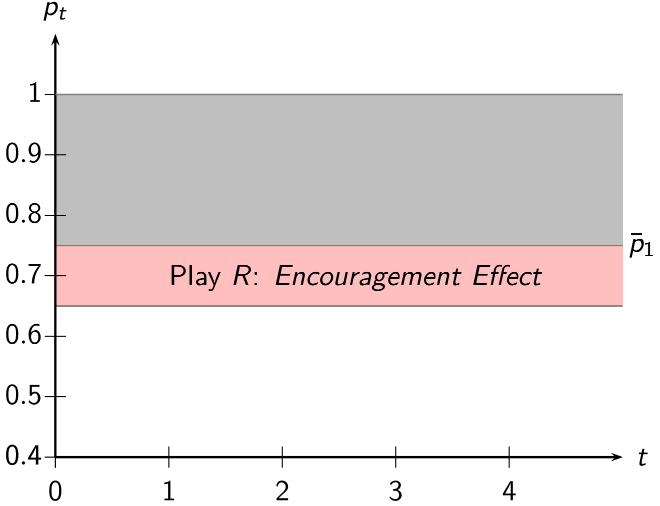
Bad News: Best-Reply



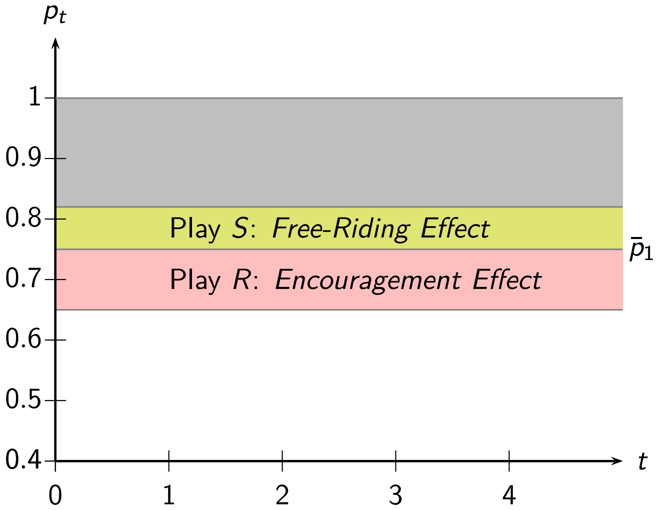
Bad News: Best-Reply



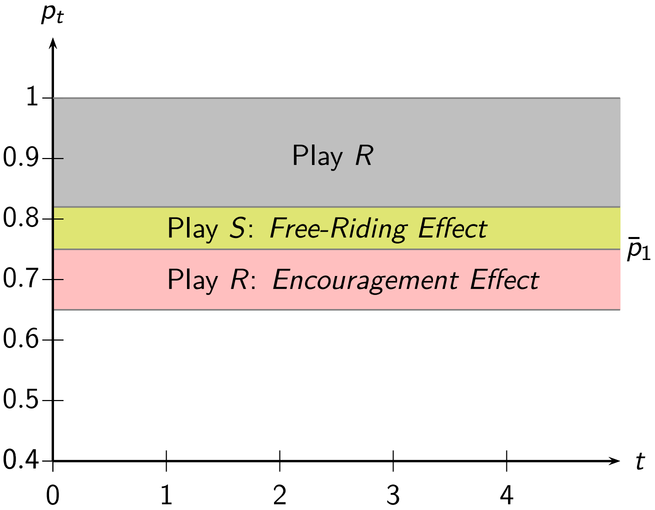
Bad News: Best-Reply



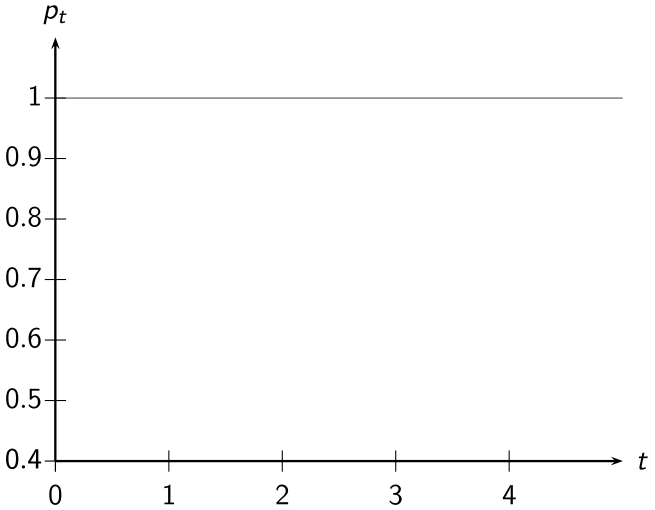
Bad News: Best-Reply



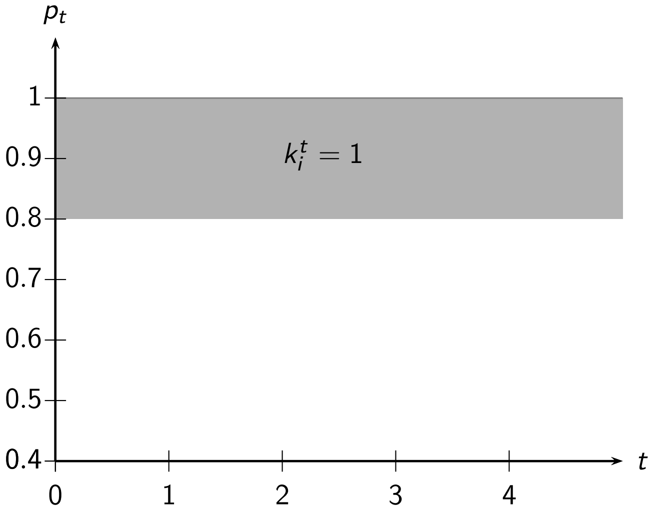
Bad News: Best-Reply



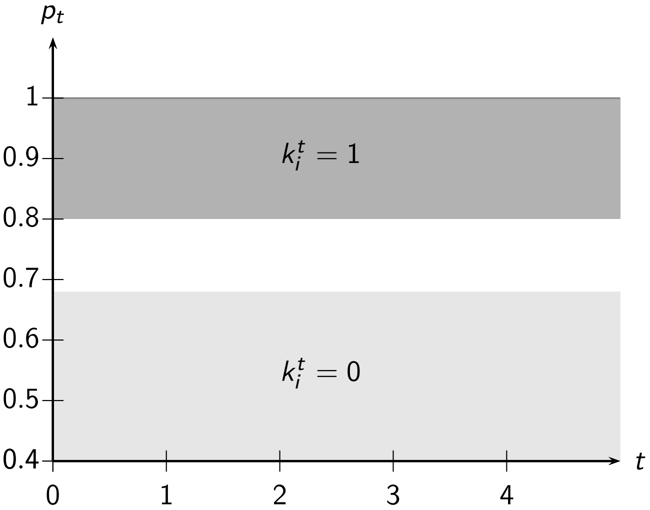
(Symmetric Markov Perfect) Equilibrium



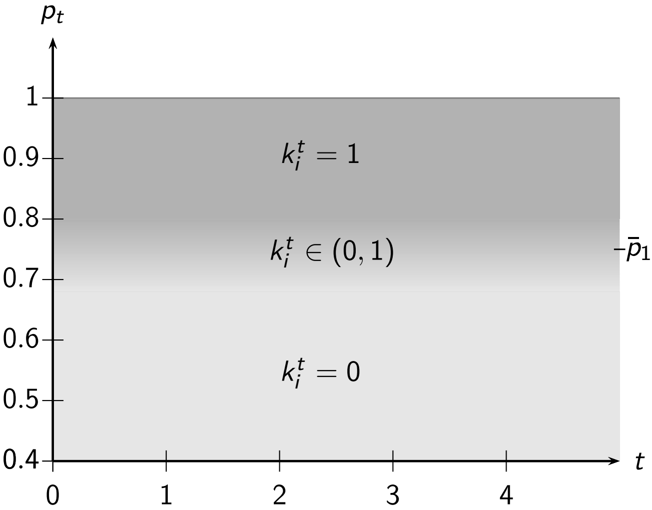
(Symmetric Markov Perfect) Equilibrium



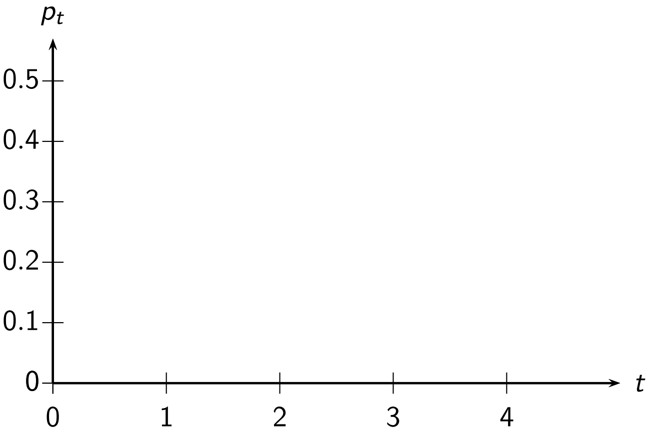
(Symmetric Markov Perfect) Equilibrium



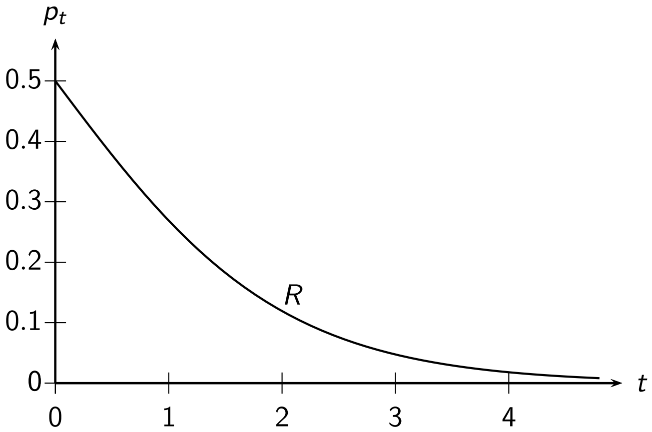
(Symmetric Markov Perfect) Equilibrium



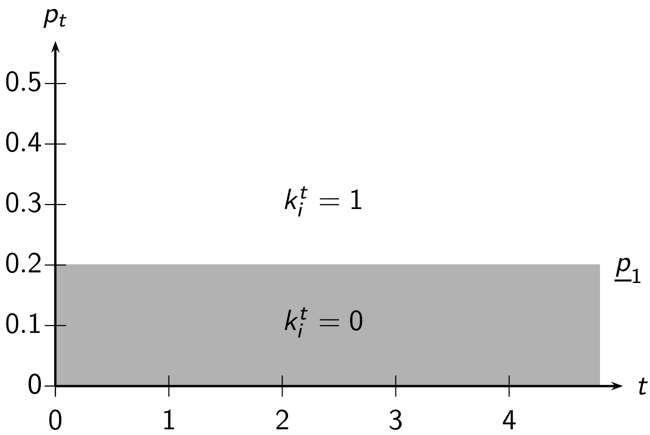
Good News: Best-Reply



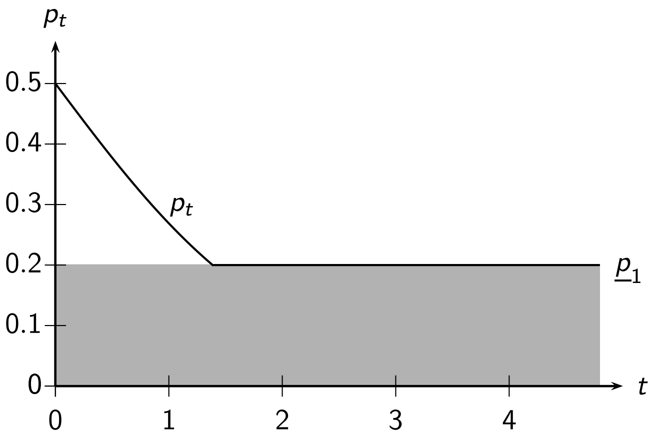
Good News: Best-Reply



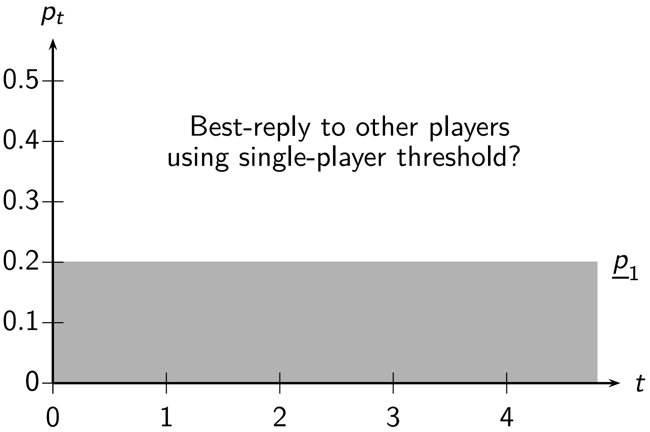
Good News: Best-Reply



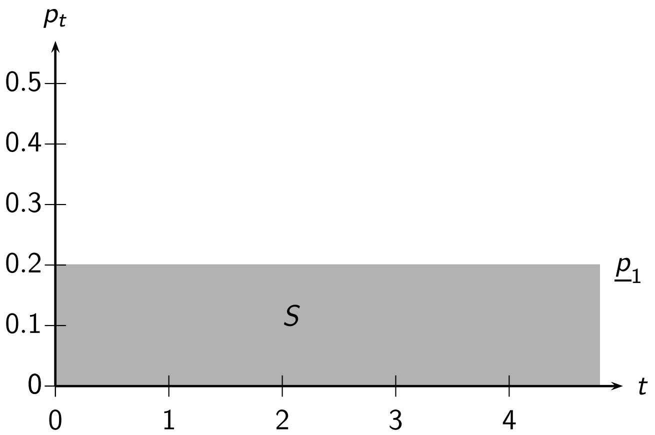
Good News: Best-Reply



Good News: Best-Reply

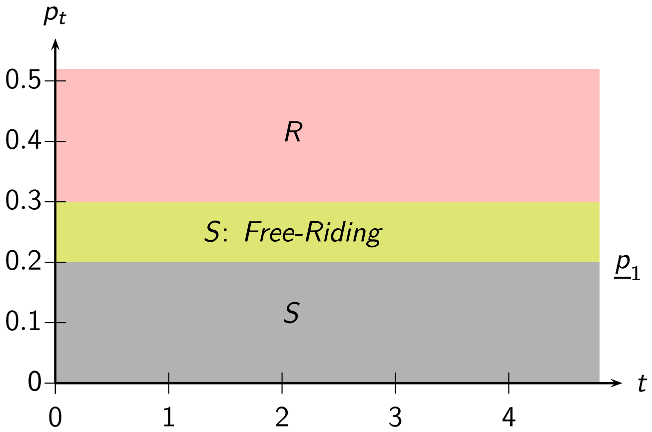


Good News: Best-Reply

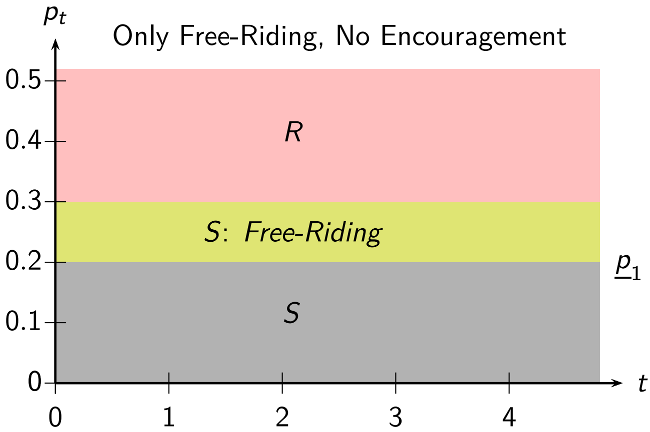


Good News: Best-Reply

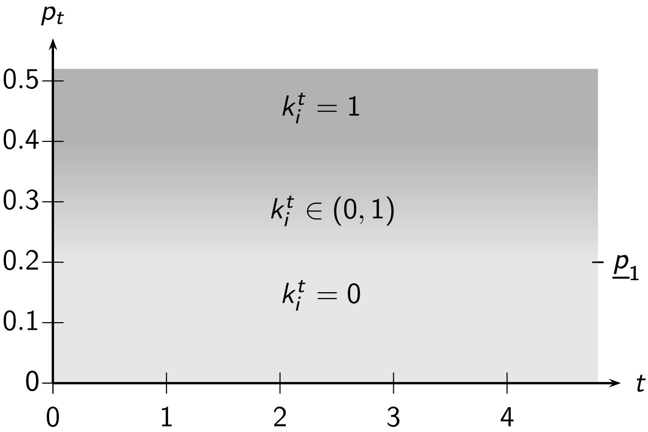
Good News: Best-Reply



Good News: Best-Reply



Equilibrium



Different Dynamics lead to different incentives:

Good News: Only Free-riding affects behavior.

Bad News: Both Free-riding and Encouragement.

Different Dynamics lead to different incentives:

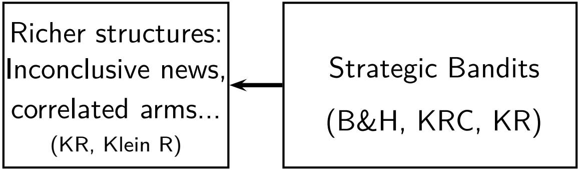
Good News: Only Free-riding affects behavior.

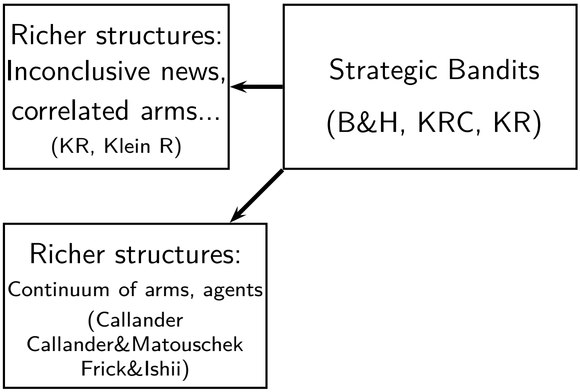
Bad News: Both Free-riding and Encouragement.

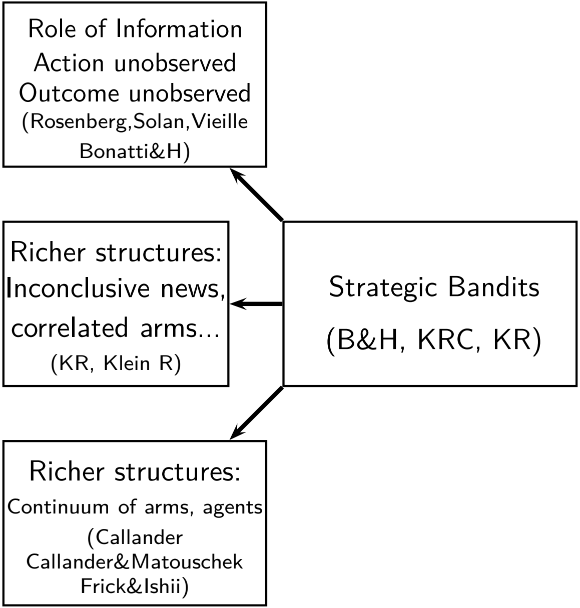
The bad news case is more representative, *e.g.*, both effects are present with inconclusive news and Brownian motion.

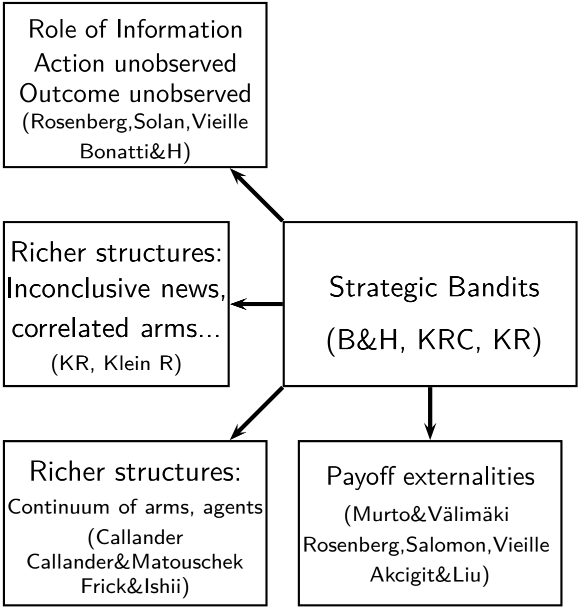
Strategic Bandits

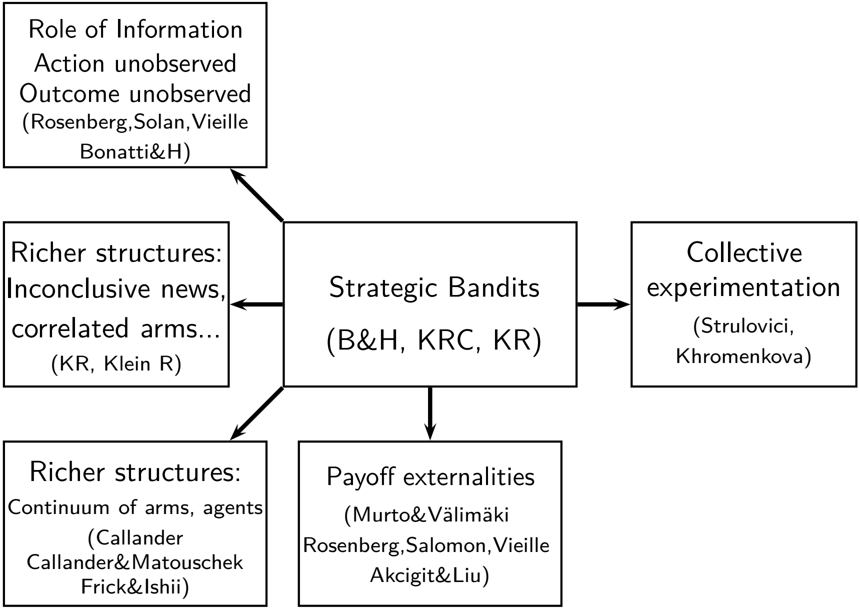
(B&H, KRC, KR)

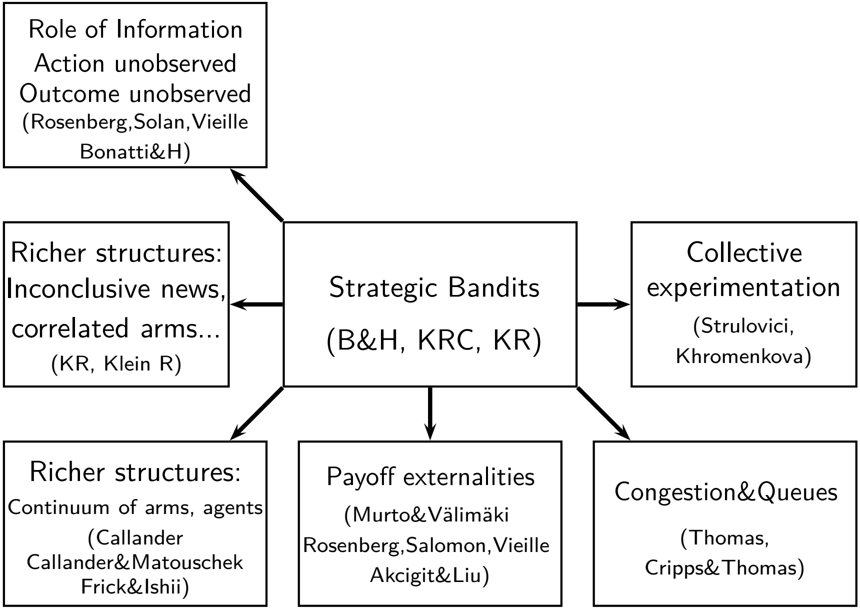

Richer structures:
Inconclusive news,
correlated arms...
(KR, Klein R)
Strategic Bandits
(B&H, KRC, KR)


Richer structures:
Inconclusive news,
correlated arms...
(KR, Klein R)
Strategic Bandits
(B&H, KRC, KR)
Richer structures:
Continuum of arms, agents
(Callander
Callander&Matouschek
Frick&Ishii)

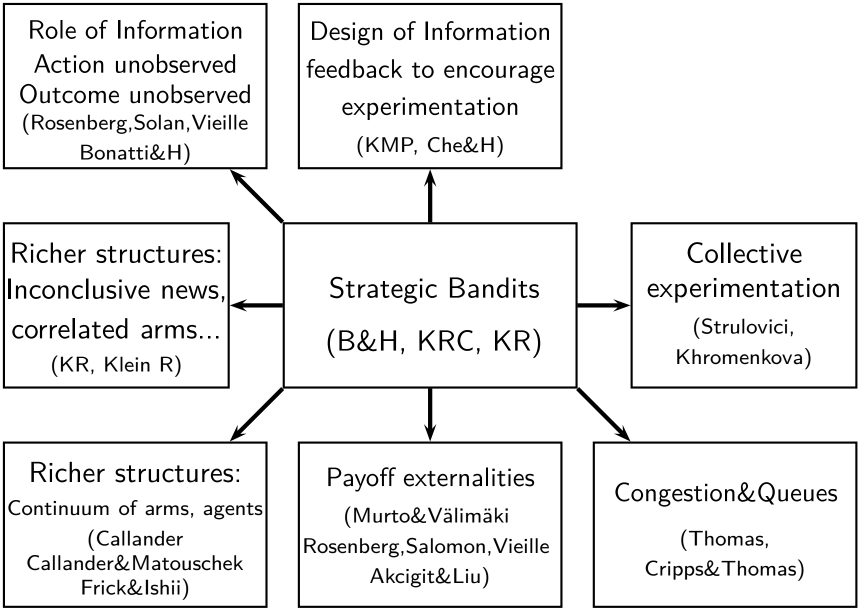

Role of Information
Action unobserved
Outcome unobserved
(Rosenberg,Solan,Vieille
Bonatti&H)
Richer structures:
Inconclusive news,
correlated arms...
(KR, Klein R)
Strategic Bandits
(B&H, KRC, KR)
Richer structures:
Continuum of arms, agents
(Callander
Callander&Matouschek
Frick&Ishii)

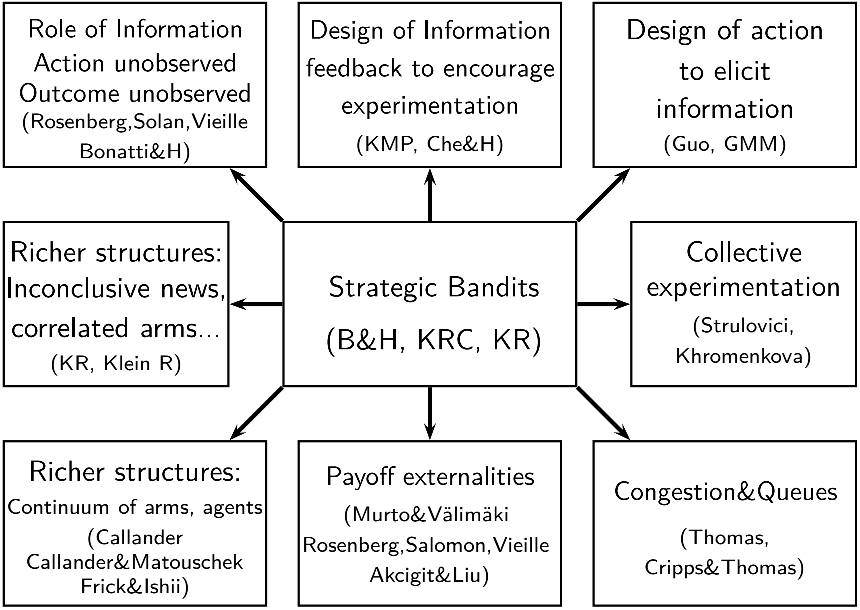

Role of Information
Action unobserved
Outcome unobserved
(Rosenberg,Solan,Vieille
Bonatti&H)
Richer structures:
Inconclusive news,
correlated arms...
(KR, Klein R)
Strategic Bandits
(B&H, KRC, KR)
Richer structures:
Continuum of arms, agents
(Callander
Callander&Matouschek
Frick&Ishii)
Payoff externalities
(Murto&Välimäki
Rosenberg,Salomon,Vieille
Akcigit&Liu)

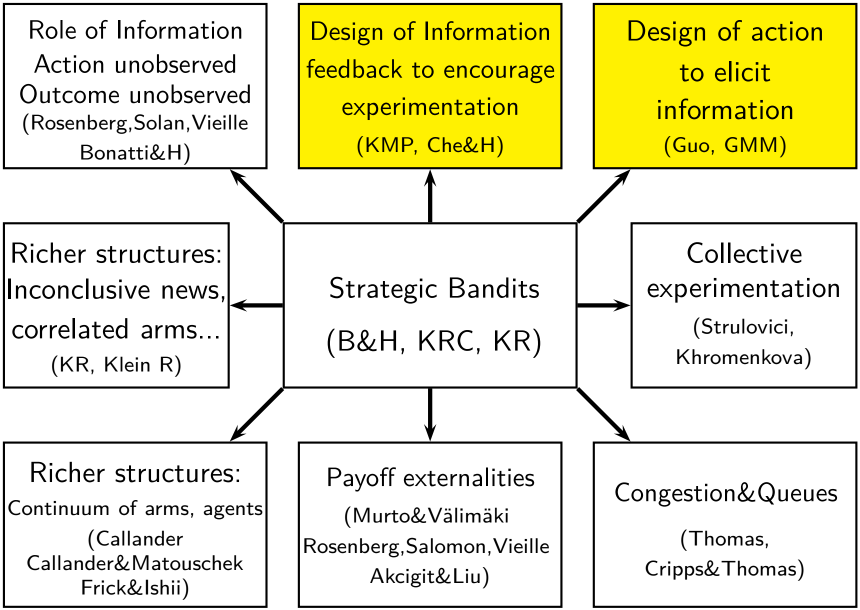

Role of Information
Action unobserved
Outcome unobserved
(Rosenberg,Solan,Vieille
Bonatti&H)
Richer structures:
Inconclusive news,
correlated arms...
(KR, Klein R)
Strategic Bandits
(B&H, KRC, KR)
Collective
experimentation
(Strulovici,
Khromenkova)
Richer structures:
Continuum of arms, agents
(Callander
Callander&Matouschek
Frick&Ishii)
Payoff externalities
(Murto&Välimäki
Rosenberg,Salomon,Vieille
Akcigit&Liu)



Role of Information
Action unobserved
Outcome unobserved
(Rosenberg,Solan,Vieille
Bonatti&H)
Richer structures:
Inconclusive news,
correlated arms...
(KR, Klein R)
Strategic Bandits
(B&H, KRC, KR)
Collective
experimentation
(Strulovici,
Khromenkova)
Richer structures:
Continuum of arms, agents
(Callander
Callander&Matouschek
Frick&Ishii)
Payoff externalities
(Murto&Välimäki
Rosenberg,Salomon,Vieille
Akcigit&Liu)
Congestion&Queues
(Thomas,
Cripps&Thomas)


Role of Information
Action unobserved
Outcome unobserved
(Rosenberg,Solan,Vieille
Bonatti&H)
Design of Information
feedback to encourage
experimentation
(KMP, Che&H)
Richer structures:
Inconclusive news,
correlated arms...
(KR, Klein R)
Strategic Bandits
(B&H, KRC, KR)
Collective
experimentation
(Strulovici,
Khromenkova)
Richer structures:
Continuum of arms, agents
(Callander
Callander&Matouschek
Frick&Ishii)
Payoff externalities
(Murto&Välimäki
Rosenberg,Salomon,Vieille
Akcigit&Liu)
Congestion&Queues
(Thomas,
Cripps&Thomas)


Role of Information
Action unobserved
Outcome unobserved
(Rosenberg,Solan,Vieille
Bonatti&H)
Design of Information
feedback to encourage
experimentation
(KMP, Che&H)
Design of action
to elicit
information
(Guo, GMM)
Richer structures:
Inconclusive news,
correlated arms...
(KR, Klein R)
Strategic Bandits
(B&H, KRC, KR)
Collective
experimentation
(Strulovici,
Khromenkova)
Richer structures:
Continuum of arms, agents
(Callander
Callander&Matouschek
Frick&Ishii)
Payoff externalities
(Murto&Välimäki
Rosenberg,Salomon,Vieille
Akcigit&Liu)
Congestion&Queues
(Thomas,
Cripps&Thomas)


Role of Information
Action unobserved
Outcome unobserved
(Rosenberg,Solan,Vieille
Bonatti&H)
Design of Information
feedback to encourage
experimentation
(KMP, Che&H)
Design of action
to elicit
information
(Guo, GMM)
Richer structures:
Inconclusive news,
correlated arms...
(KR, Klein R)
Strategic Bandits
(B&H, KRC, KR)
Collective
experimentation
(Strulovici,
Khromenkova)
Richer structures:
Continuum of arms, agents
(Callander
Callander&Matouschek
Frick&Ishii)
Payoff externalities
(Murto&Välimäki
Rosenberg,Salomon,Vieille
Akcigit&Liu)
Congestion&Queues
(Thomas,
Cripps&Thomas)

Information Design

Kremer, Mansour and Perry (2014)
Che and Hörner (2015)

Sequence of agents choosing between two routes ($k = S, R$).

Sequence of agents choosing between two routes ($k = S, R$).

Payoffs are common but unknown, with $\pi_k \sim \mathcal{U}[a_k, b_k]$.

Sequence of agents choosing between two routes ($k = S, R$).

Payoffs are common but unknown, with $\pi_k \sim \mathcal{U}[a_k, b_k]$.

Once route k is tried once, π_k is revealed.

Sequence of agents choosing between two routes ($k = S, R$).

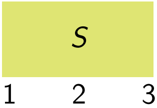
Payoffs are common but unknown, with $\pi_k \sim \mathcal{U}[a_k, b_k]$.

Once route k is tried once, π_k is revealed.

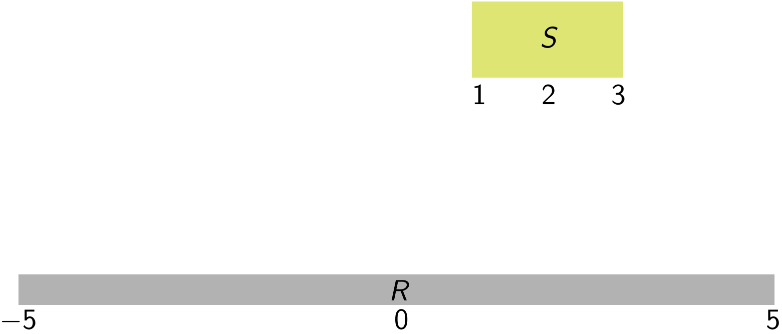
Patient, benevolent designer would like both routes to be tried.

She controls what information about earlier choices to disclose.

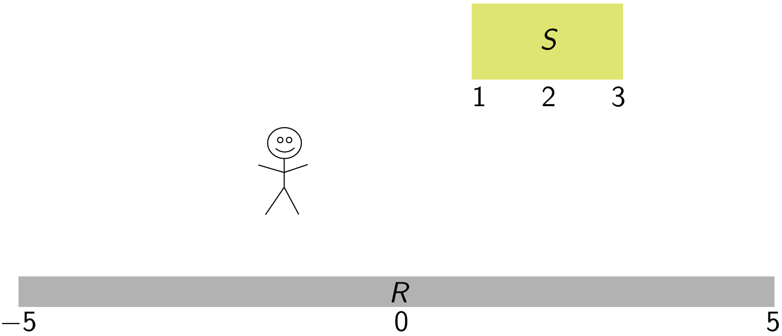
No Designer Around



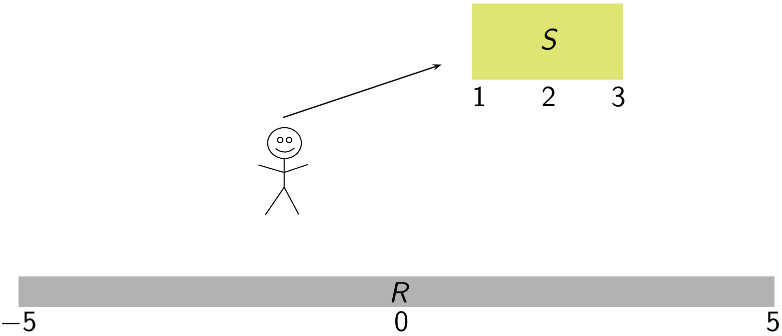
No Designer Around



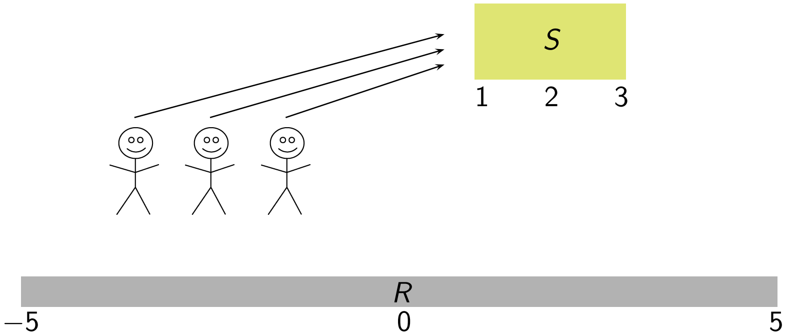
No Designer Around



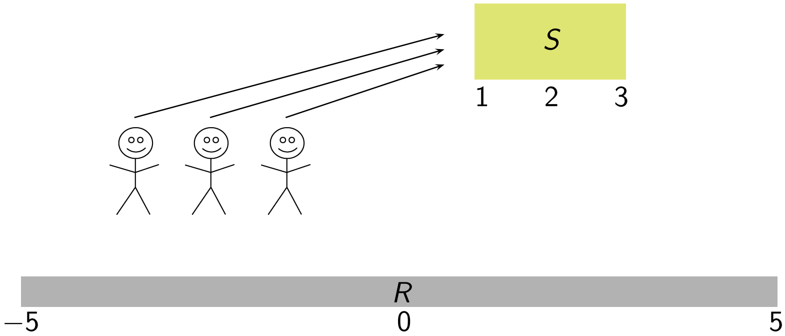
No Designer Around



No Designer Around

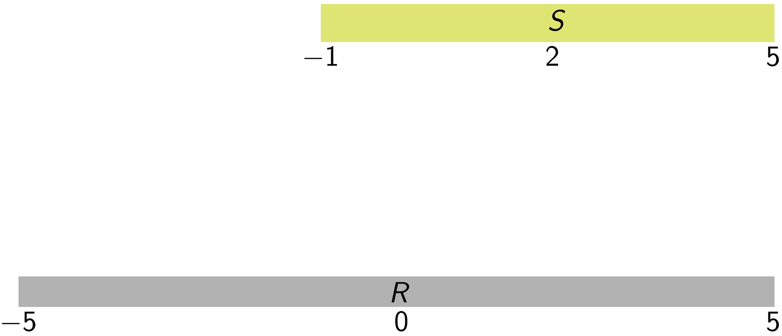


No Designer Around

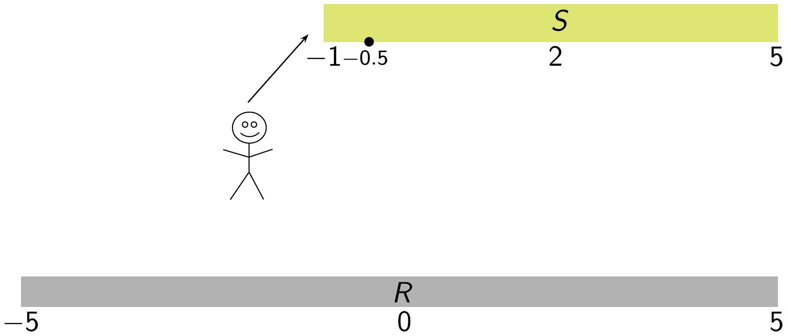


Experimentation cannot take off if $\mathbf{E}[\pi_R] < a_S$.

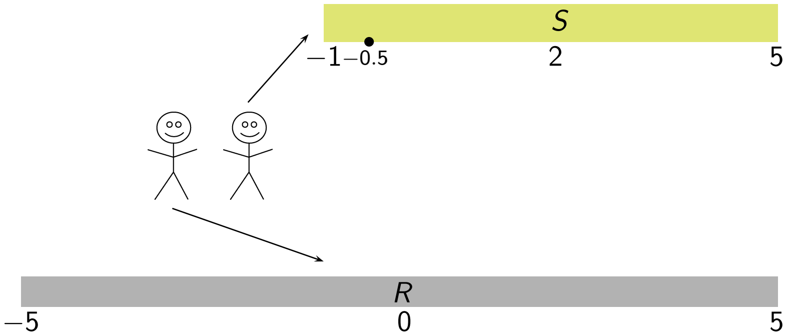
No Designer Around



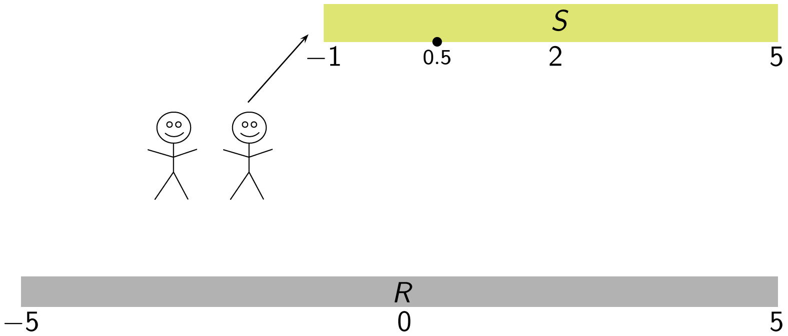
No Designer Around



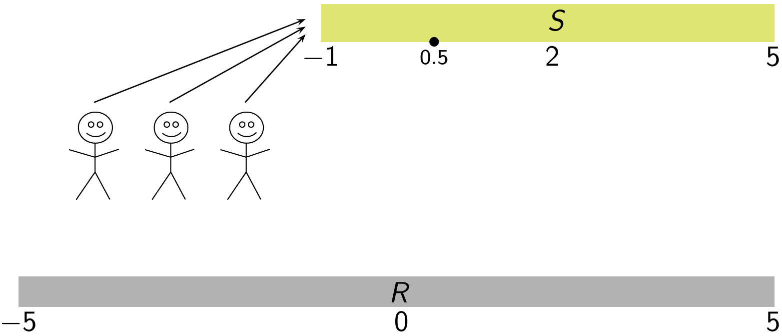
No Designer Around



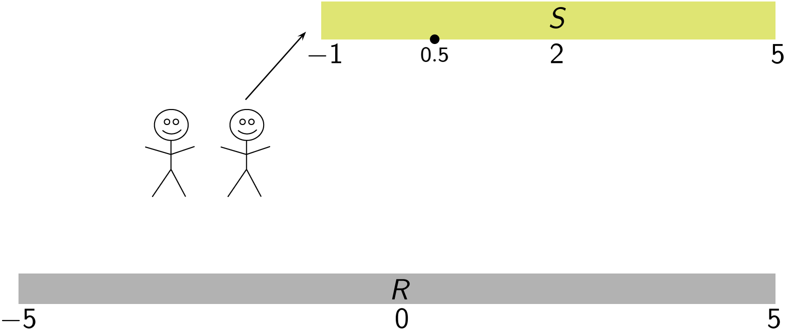
No Designer Around



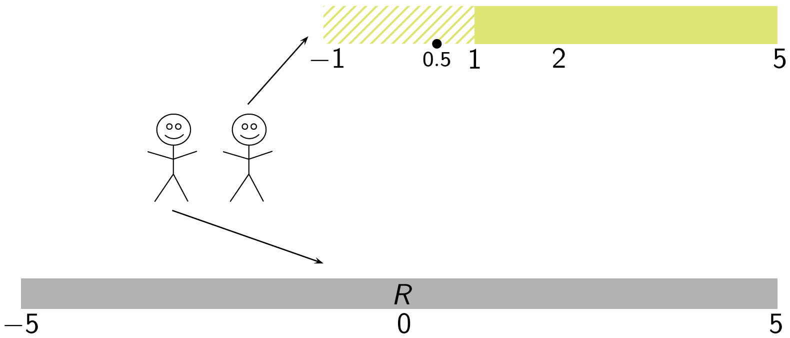
No Designer Around



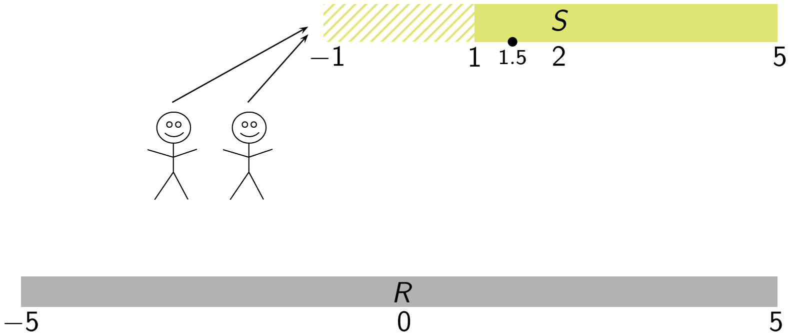
With a Designer



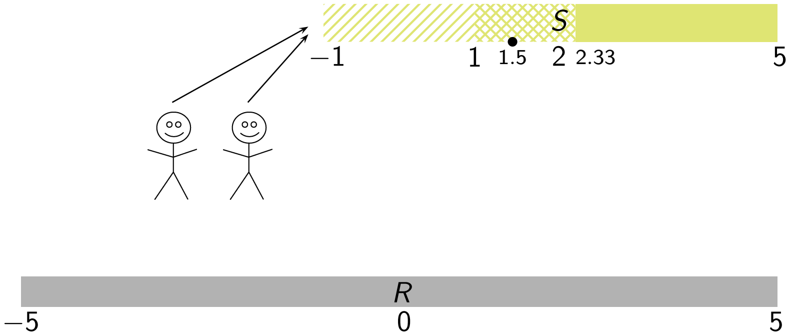
With a Designer



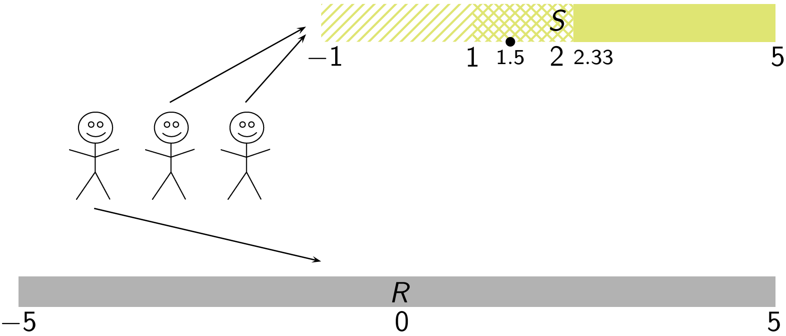
With a Designer



With a Designer



With a Designer



Experimentation proceeds if only partial information is given away.

Either it cannot get started, or it can be completed (if so desired).

Optimal structure is partitional and "bottom up."

Early agents experiment, later agents benefit.

In KMP, the uncertainty about the value is diffuse, but the number of experiments it takes to find out is very limited.

In KMP, the uncertainty about the value is diffuse, but the number of experiments it takes to find out is very limited.

Consider the opposite: the risky arm has binary value $\omega = 0, 1$.

In KMP, the uncertainty about the value is diffuse, but the number of experiments it takes to find out is very limited.

Consider the opposite: the risky arm has binary value $\omega = 0, 1$.

But it takes time to learn –few agents bother with feedback.

Specifically, suppose that only satisfied agents report, if at all.

If (endogenous) fraction μ_t tries out over interval of length $\mathrm{d}t$,

$$\text{Prob(positive feedback)} = \lambda \cdot \mu_t \cdot \omega \mathrm{d}t,$$

In KMP, the uncertainty about the value is diffuse, but the number of experiments it takes to find out is very limited.

Consider the opposite: the risky arm has binary value $\omega = 0, 1$.

But it takes time to learn –few agents bother with feedback.

Specifically, suppose that only satisfied agents report, if at all.

If (endogenous) fraction μ_t tries out over interval of length $\mathrm{d}t$,

$$\text{Prob(positive feedback)} = \lambda \cdot \mu_t \cdot \omega \mathrm{d}t,$$

In KMP, the uncertainty about the value is diffuse, but the number of experiments it takes to find out is very limited.

Consider the opposite: the risky arm has binary value $\omega = 0, 1$.

But it takes time to learn –few agents bother with feedback.

Specifically, suppose that only satisfied agents report, if at all.

If (endogenous) fraction μ_t tries out over interval of length $\mathrm{d}t$,

$$\text{Prob(positive feedback)} = \lambda \cdot \mu_t \cdot \omega \mathrm{d}t,$$

There are two kinds of agents:

A fraction ρ an opportunity cost of $c \in (p_0, 1)$;

The others experiment anyhow.

There are two kinds of agents:

A fraction ρ an opportunity cost of $c \in (p_0, 1)$;

The others experiment anyhow.

Because $c > p_0$, learning starts only with these others.
Yet quicker learning might be socially desirable.

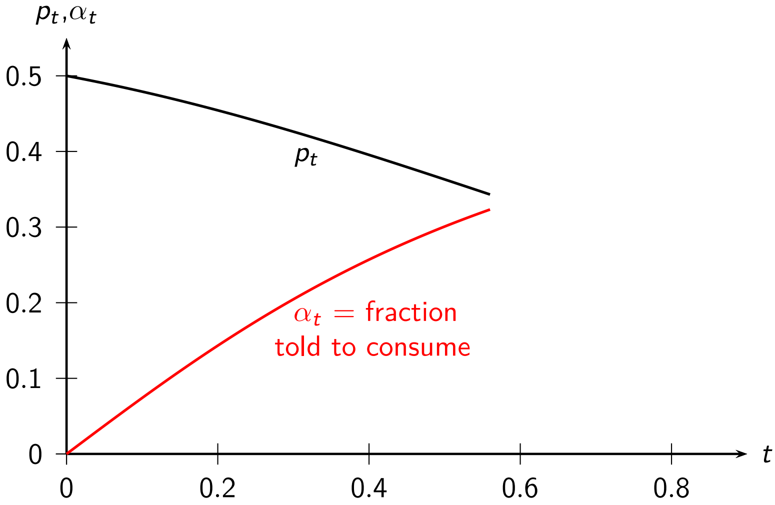
There are two kinds of agents:

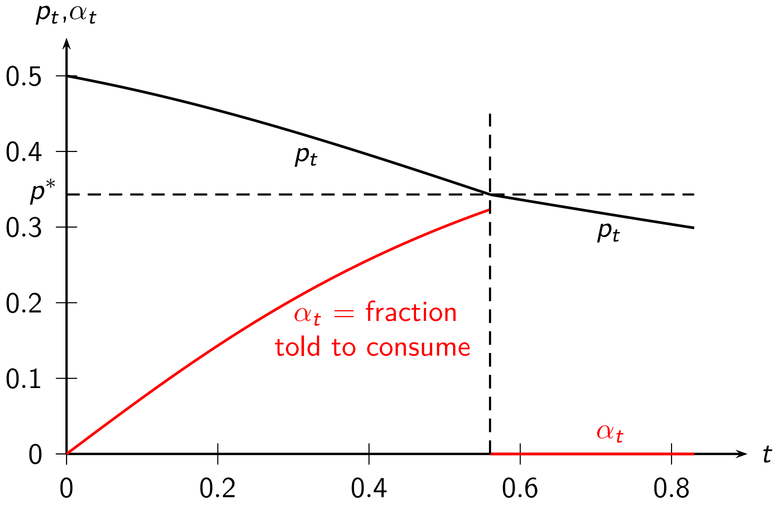
- A fraction ρ an opportunity cost of $c \in (p_0, 1)$;
- The others experiment anyhow.

Because $c > p_0$, learning starts only with these others.
Yet quicker learning might be socially desirable.

How can the social planner incentivize experimentation?

- Tell everyone to experiment when she has learnt $\omega = 1$.
- Tell a random fraction of agents to do so when she hasn't yet.


p_t,α_t
0.5
0.4
0.3
0.2
0.1
0
p_t
α_t = fraction
told to consume
0
0.2
0.4
0.6
0.8
t



p_t, α_t
0.5
0.4
p^*
0.3
0.2
0.1
0
p_t
p_t
α_t = fraction
told to consume
α_t
0
0.2
0.4
0.6
0.8
t

Credibility takes time to build: early recommendations must be truthful.

Late recommendations must be truthful as well: “spamming” is single-peaked.

Credibility takes time to build: early recommendations must be truthful.

Late recommendations must be truthful as well: “spamming” is single-peaked.

Ultimately, as in KMP, experimentation amount is first-best, but the price to pay is costly delay.

Delegation

Guo (2014)
Grenadier, Malenko and Malenko (2015)

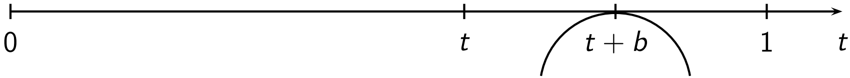
When agent has additional private information, we don't only want to control what he sees, but elicit his information.

Dynamics is a source of commitment. Rather than start with a bandit, consider the following version of Crawford-Sobel:

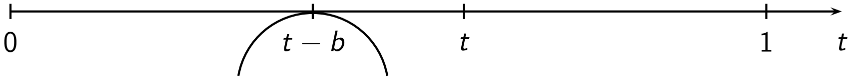
State $t \sim \mathcal{U}[0,1]$.
Action $y \in [0,1]$.
Receiver's utility $-(y-t)^2$.
Sender's utility $-(y-t \pm b)^2$, $b \in (0, 1/2)$.

Twist: the action is the time at which to act; players can wait.

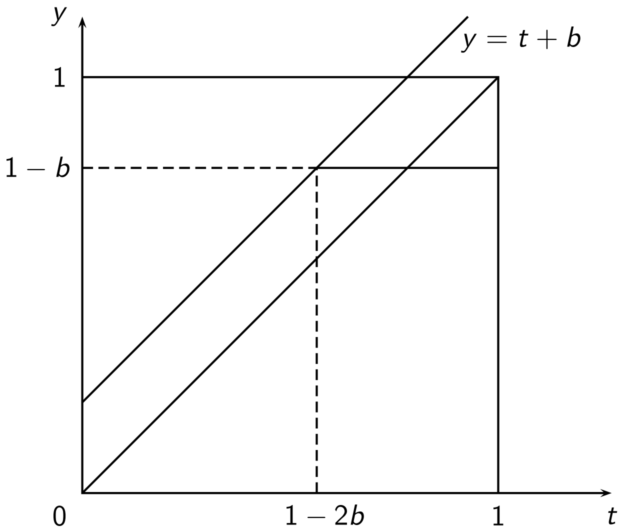
Positive bias (Preference for delay)



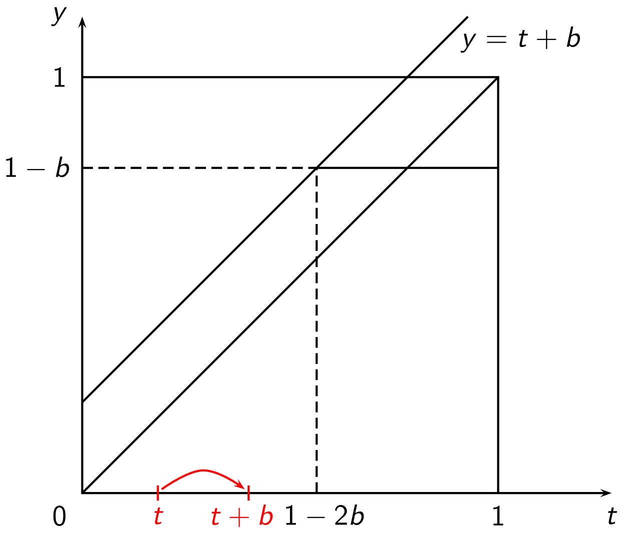
Negative bias (Preference for early action)



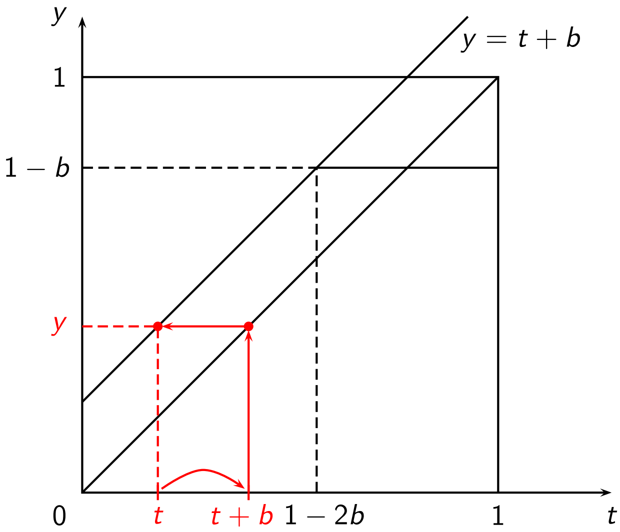
Positive Bias



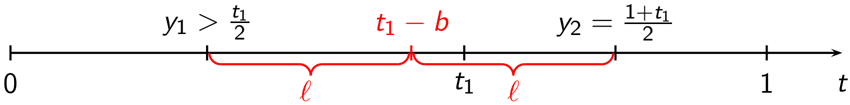
Positive Bias



Positive Bias



Negative Bias



In Guo (2015):

Biased agent has better information about risky arm's type.

Principal designs the optimal mechanism.

Implementable as a delegation mechanism.

In Guo (2015):

Biased agent has better information about risky arm's type.

Principal designs the optimal mechanism.

Implementable as a delegation mechanism.

In Grenadier, Malenko & Malenko (2015):

No learning/incomplete information.

Principal faces standard "real options" investment problem.

Agent privately knows a factor multiplying the value.

In Guo (2015):

Biased agent has better information about risky arm's type.
Principal designs the optimal mechanism.
Implementable as a delegation mechanism.

In Grenadier, Malenko & Malenko (2015):

No learning/incomplete information.
Principal faces standard "real options" investment problem.
Agent privately knows a factor multiplying the value.

But many common conclusions, in particular:
no commitment required iff preference for early switch.

As mentioned, lots of topics I have omitted here.

Design of information channel: Ely, Renault Solan Vieille,...

As mentioned, lots of topics I have omitted here.

- Design of information channel: Ely, Renault Solan Vieille,...

But also lots of exciting open questions. Need:

- Tools/principles to “replace” the “Myersonian” toolbox in the absence of transfers.
- Stronger interplay with applied work:
  Development, IO, trade, economics of organizations, etc.