Strategic Feedback Mechanisms and Experimentation

Attila Ambrus
Duke

Luke Zhao
Duke

Runxin Zhi
Duke

May 2026

INTRODUCTION

- Experimentation and adoption of new technologies/products intensely studied topic both theoretically and empirically.
- Individual and social incentives typically do not align — starting experimentation might not be individually rational even when has positive social value.
- Question: how incentives to experiment can be influenced.
- One way is strategically providing information. Kremer, Mansour and Perry (2014, JPE) and Che and Horner (2018, QJE) investigate how to optimally provide extra information a designer has relative to information learned from consumers using the product. Cold start problem: experimentation can only be started gradually.

RESEARCH QUESTION

- As opposed to previous work in which consumers know whether they are early or late adaptors, we focus on settings in which they do not have this information.
- Example: a physician experimenting with a non FDA-approved medical substance. Decides whether to recommend a patient taking the substance. If a patient takes the substance, the doctor receives a signal about its effectiveness. A consumer, when receives a recommendation, does not know how much and what feedback the doctor received beforehand.
- Other examples: personalized *Amazon* recommendations, *AliExpress* emails.
- We take an information design approach: the designer can commit to a recommendation policy contingent on history of signals observed.

PREVIEW OF THE RESULTS

- If consumers do not know the product's available time and purchasing history, welfare-improving experimentation is possible even if it is well-known that the designer does not hold extra information except what learned from past purchases.
- Under general consumer signals, the designer can implement his first best if the designer maximizes intertemporal expected consumer surplus.
- While the problem is not stationary, the designer always has a threshold-form second best policy with possible mixing only at the threshold.
- We provide full characterizations of the second best policies under perfect good news and perfect bad news feedback signal structures, as well as general preferences of the designer.

RELATED LITERATURES

- *Incentivizing Exploration with Information*: Che and Hörner (2018); Kremer et al. (2014); Madsen and Shmaya (2025); Lyu (2026).
  - *Empirical*: Hanna et al. (2014); Ayalew et al. (2022).
  - *In CS*: Immorlica et al. (2018); Mansour et al. (2022); Slivkins (2023) and Simchowitz and Slivkins (2024).

- *Strategic Experimentation*: Bolton and Harris (1999); Keller et al. (2005); Keller and Rady (2010; 2015).

- *Information Design*: Kamenica and Gentzkow (2011); Bergemann and Morris (2019).

- *Social Learning and Information Cascade*: Banerjee (1992); Bikhchandani et al. (2024)

A Motivational Example

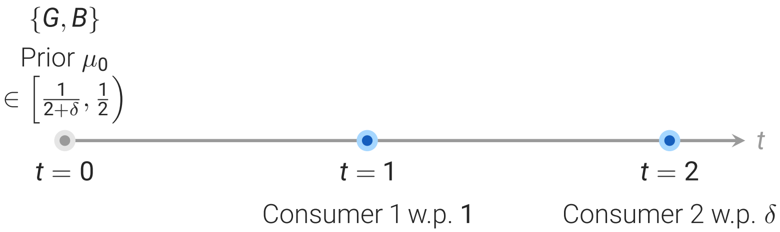
A TWO-PERIOD EXAMPLE

We start from a simple example with the following simplifying assumptions:

- There are at most two consumers.
- The designer receives perfect feedback signals from the consumers.

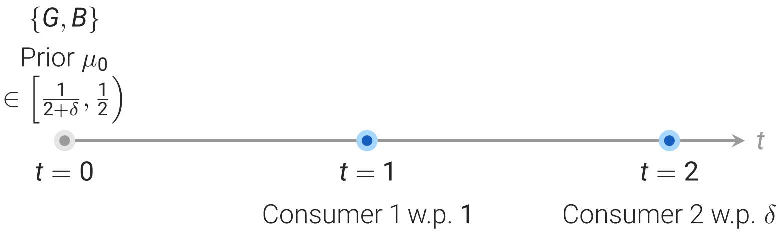
These are just for simplicity and will be relaxed in the formal analysis later.

A SIMPLE EXAMPLE



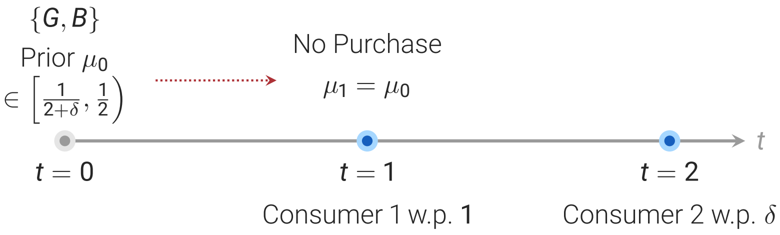
- A designer recommends a product to (at most) 2 consumers.
- Quality is initially unknown to everyone but is revealed to the designer after a single purchase.

A SIMPLE EXAMPLE



- Payoff u_t of consumer t: $\{\text{Buy } G, \text{No purchase}, \text{Buy } B\} = \{1, 0, -1\}$.
- The designer commits to a recommendation plan that maximizes the ex-ante expected intertemporal consumer surplus $\mathbb{E}[u_1 + \delta u_2]$.

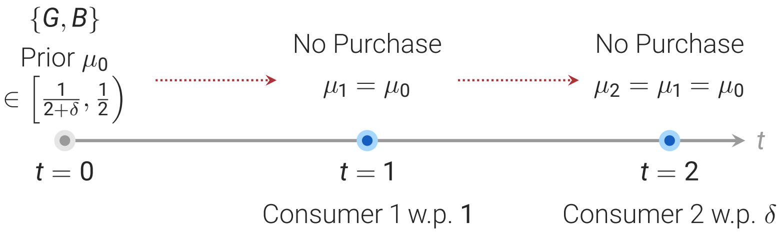
A SIMPLE EXAMPLE



If consumers know their arrival time:

- Consumer 1 will ignore any recommendation and choose not to purchase.

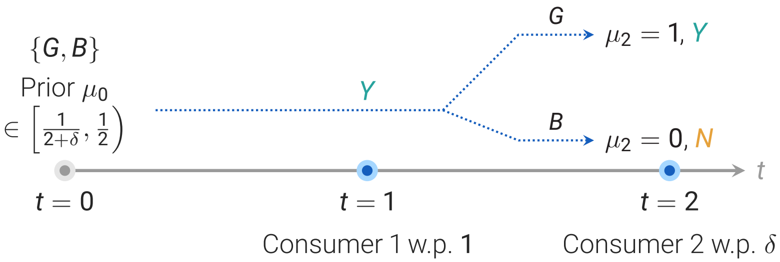
A SIMPLE EXAMPLE



If consumers know their arrival time:

- Consumer 1 will ignore any recommendation and choose not to purchase.
- No update possible: Consumer 2 will not purchase either.
- Designer's payoff (consumer surplus) $= 0$.

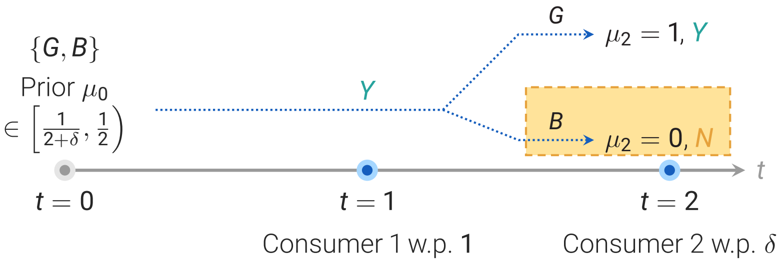
A SIMPLE EXAMPLE



If consumers do not know their arrival time, consider the following policy:

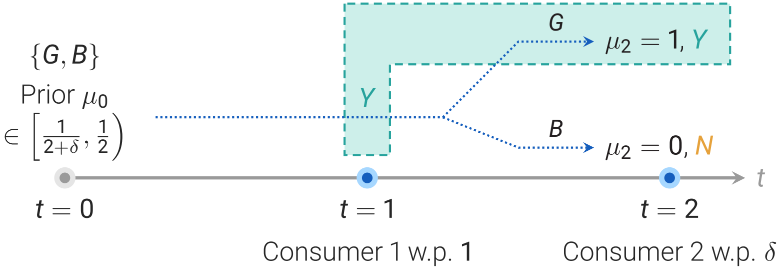
- The designer recommends purchasing (Y) to Consumer 1.
- If the product is G, recommend Y to Consumer 2. Otherwise, recommend not purchasing (N) to Consumer 2.

INCENTIVE COMPATIBLE?



- A consumer's posterior belief after N is 0.

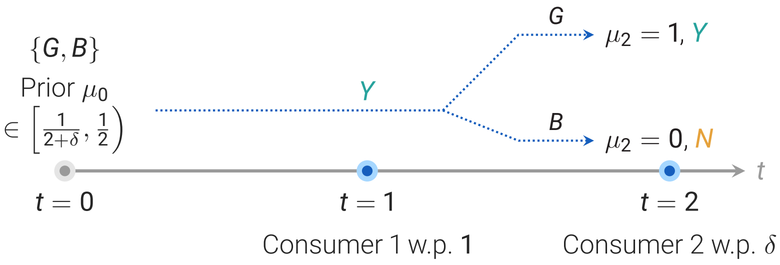
INCENTIVE COMPATIBLE?



- A consumer's posterior belief after N is 0.
- A consumer's posterior belief after Y:

$$\frac{1}{1+\delta\mu_0} \times \mu_0 + \frac{\delta\mu_0}{1+\delta\mu_0} \times 1 \qquad \text{IC if posterior} \geqslant \frac{1}{2} \text{ or } \mu_0 \geqslant \frac{1}{2+\delta}.$$

DESIGNER'S PAYOFF



The designer's payoff (consumer surplus) is positive if IC:

$$\mu_0 \times 1 + (1 - \mu_0) \times (-1) + \delta\mu_0 \times 1 = (2 + \delta)\mu_0 - 1.$$

A SIMPLE EXAMPLE: TAKEAWAYS

- The designer can facilitate experimentation and increase welfare even without additional information.
- For consumers, IC is maintained in an intertemporal way: earlier losses from *exploration* are compensated by later benefits from *exploitation*.

The Model

THE MODEL

- Discrete time, infinite horizon: $t = 1, 2, \cdots$
- A designer (he) faces a sequence of consumers (she), at most 1 at each period.
- The designer is long-lived, while consumers are myopic.
- The designer and all consumers have a common prior μ_0.
- After each purchase, consumer receives signal about product quality according to $\pi : \{G, B\} \to \Delta(S)$. Leaves feedback to the system.
- The designer commits to a (history-dependent) recommendation policy at the beginning of the game.
  - We will without loss consider direct mechanisms: $r : H \to \Delta(\{Y, N\})$.

CONSUMERS

- Total number of consumers arriving is random.
  - First consumer always arrives.
  - If consumer at t arrives, with probability δ consumer at $t+1$ also arrives. With probability $1-\delta$ sequence stops, no more arrivals.
- Payoffs: for every consumer:

$$u_t = \begin{cases} 1 & \text{if consuming } G \text{ product,} \\ 0 & \text{if not consuming,} \\ -1 & \text{if consuming } B \text{ product.} \end{cases}$$

THE DESIGNER'S PAYOFF

$$\sum_{t=1}^{\infty} \delta^{t-1} \mathbb{E}(v_t), \quad \text{where} \quad v_t = \begin{cases} 1 & \text{if consuming } G \text{ product at } t, \\ 0 & \text{if not consuming at } t, \\ a & \text{if consuming } B \text{ product at } t. \end{cases}$$

- $a = -1$: The designer maximizes intertemporal expected consumer surplus.
- $a > -1$: The designer would like to *over-experiment*.

  $a = 1$: The designer is a profit-maximizing seller.
- $a < -1$: The designer would like to *under-experiment*.
- Consumers may not internalize externalities from consumption.

The First Best

THE FIRST BEST PROBLEM

- Ignore incentive compatibility constraints in the first best problem.
- The first best problem is a stationary Markov decision problem with state μ.
- There exists an optimal policy that is Markov. Solve using Bellman:

$$W(\mu) = \max\{\delta W(\mu), (1 - a)\mu + a + \delta \mathbb{E}_\pi[W(\mu') \mid \mu]\}.$$

THRESHOLD POLICY

Theorem.
There exists a threshold $\underline{\mu} \in (0, 1)$ such that the designer recommends *Y* for sure if $\mu > \underline{\mu}$ and *N* for sure if $\mu < \underline{\mu}$.

- A standard argument showing that W is increasing, and strictly increasing if $W(\cdot) > 0$.
- $\underline{\mu} < 1/2$ if $a = -1$ due to experimentation incentives.

The Second Best

THE SECOND BEST POLICIES

We will show that:

- If the designer's goal is to *maximize intertemporal consumer surplus* (i.e., $a = -1$), the first best is incentive compatible.
- If the designer would like to *over- or under- experiment* (i.e., $a > -1$ or $a < -1$), the second best maintains a threshold form and in general involves mixing only at the threshold belief.

CONSUMERS' INCENTIVE COMPATIBILITY

- Since consumers do not observe their arrival time, they only update beliefs using the recommendation $r \in \{Y, N\}$.
- Every consumer recommended Y shares the same μ^Y, and every consumer recommended N shares the same μ^N.
- ICC-Y: $\mu^Y \geqslant 1/2$ and ICC-N: $\mu^N \leqslant 1/2$.

$$\mu^Y = \sum_t \mathbb{P}(Y \text{ at time } t \mid Y)\mathbb{E}(\mu_t \mid Y \text{ at time } t)$$

The IC constraints are *global* constraints.

CONSUMERS' INCENTIVE COMPATIBILITY

Aggregate by time vs. Aggregate by beliefs:

$$\mu^Y = \sum_t \mathbb{P}(Y \text{ at time } t \mid Y)\mathbb{E}(\mu_t \mid Y \text{ at time } t)$$
$$= \sum_\mu \frac{\text{Measure of } Y \text{ at } \mu}{\text{Overall measure of } Y} \cdot \mu.$$

- This measure contains probability of arrival, probability of the signal realization, and probability of recommendation according to the committed policy.

OCCUPANCY MEASURES

To check the incentive compatibility of the first best policy, we formalize

$$\sum_{\mu} \frac{\text{Measure of } Y \text{ at } \mu}{\text{Overall measure of } Y} \cdot \mu$$

using occupancy measures under the first best policy:

- $x(\mu)$: the expected number of periods the belief at the beginning of the period is μ, and there is a consumer arriving.
- $y(\mu)$: the expected number of periods for which the belief at the beginning of the period is μ, there is a consumer arriving, and Y is recommended.

LAW OF MOTIONS

Let $r(\cdot)$ be the first best policy. $r(\mu) = 1$ if $\mu > \underline{\mu}$, and $r(\mu) = 0$ if $\mu < \underline{\mu}$.

$$
\begin{aligned}
y(\mu) &= r(\mu)x(\mu), \\
x(\mu) &= \mathbb{I}(\mu = \mu_0) + \delta \int y(\hat{\mu}) \, \mathrm{d}\, \pi(\mu \mid \hat{\mu}) + \delta[1 - r(\mu)]x(\mu).
\end{aligned}
$$

Consider $x(\mu)$:

- The first term: Consumer 1 always arrives.
- The second term: Arrive at belief μ by experimenting at some $\hat{\mu}$.
- The third term: Arrive at belief μ by not experimenting at μ.

ICC-Y WITH OCCUPANCY MEASURE

ICC-Y is

$$\mu^Y = \int \frac{y(\mu)}{\int y(\hat{\mu})\,\mathrm{d}\,\hat{\mu}} \cdot \mu\,\mathrm{d}\,\mu \geqslant \frac{1}{2},$$

or equivalently,

$$\int y(\mu)(2\mu - 1)\,\mathrm{d}\,\mu \geqslant 0.$$

INTERTEMPORAL CONSUMER SURPLUS MAXIMIZING DESIGNER

We now consider a designer who maximizes the intertemporal consumer surplus.

Theorem.
The designer's first best policy is incentive-compatible if $a = -1$.

- ICC-N is satisfied since $\underline{\mu} < 1/2$ in the first best.
- For ICC-Y, notice that
$$W(\mu_0) = \int y(\mu)(2\mu - 1)\,\mathrm{d}\,\mu,$$
and Y is ever recommended if $W(\mu_0) \geqslant 0$.

INTERTEMPORAL CONSUMER SURPLUS MAXIMIZING DESIGNER

We now consider a designer who maximizes the intertemporal consumer surplus.

Theorem.
The designer's first best policy is incentive-compatible if $a = -1$.

- Time-blind consumers' payoffs are the weighted sum of stage payoffs, weighted by the probability of their location in the game tree.
- This makes them weight the stage payoffs the same as the designer ex ante.
- When $a = -1$, ICC-Y shares the same sign as the designer's first-best ex ante payoff.

OVER- OR UNDER-EXPERIMENT DESIGNERS

Proposition.

- If $a > -1$, the designer's second best policy is either the first best policy, or a policy while ICC-Y is binding, and ICC-N is slack.
- If $a < -1$, the designer's second best policy is either the first best policy, or a policy while ICC-N is binding, and ICC-Y is slack.

- If a is sufficiently close to -1, the first best policy is still implementable.
- Otherwise, a policy with non-binding IC can always be improved by mixing this policy with the first best policy at the beginning of the game.

THE SECOND BEST POLICIES UNDER GENERAL SIGNAL STRUCTURES

Theorem. (Threshold Form with Mixing)
For $a \in (-\infty, 1]$ and a general signal structure π, there exists a second best policy that is a threshold policy, with possible mixing only at the threshold belief.

- The second best problem can be transformed to satisfy standard Markov conditions with an auxiliary designer ($a = -1$), leading to a threshold policy. Mixing is in general necessary when IC is binding. [Proof Sketch]
- Both the designer and the consumers prefer experimentation at higher beliefs, given the binary state space and a stationary signal structure.

Perfect Good News

PERFECT GOOD NEWS

Random good news g under G. No news n otherwise.

$$\begin{bmatrix} \mathbb{P}(g \mid G) & \mathbb{P}(n \mid G) \\ \mathbb{P}(g \mid B) & \mathbb{P}(n \mid B) \end{bmatrix} = \begin{bmatrix} \alpha & 1-\alpha \\ 0 & 1 \end{bmatrix}.$$

BELIEFS AND OCCUPANCY MEASURES

- The belief absent of good news is monotonically decreasing.
- μ_n: the belief after $n-1$ experiments and no good news yet.
- The simple belief evolution allows a linear-programming formulation of the designer's problem using occupancy measures:

$$
\begin{aligned}
x_1 &= 1 + \delta(x_1 - y_1), \\
x_{n+1} &= \delta(1-\alpha\mu_n)y_n + \delta(x_{n+1} - y_{n+1}), \quad n > 1, \\
x_G &= \delta x_G + \delta \sum_{n \geqslant 1} \alpha \mu_n y_n.
\end{aligned}
$$

THE DESIGNER'S PROBLEM

$$
\begin{aligned}
\max_{\{(x_n, y_n)\}} \quad & \sum_{n \geqslant 1} [\mu_n + a(1-\mu_n)] y_n + x_G, \\
\text{s to.} \quad & \frac{\sum_{n \geqslant 1} \mu_n y_n + 1 \cdot x_G}{\sum_{n \geqslant 1} y_n + x_G} \geqslant \frac{1}{2}, && \text{(ICC-Y)} \\
& \frac{\sum_{n \geqslant 1} (x_n - y_n) \mu_n}{\sum_{n \geqslant 1} (x_n - y_n)} \leqslant \frac{1}{2}, , && \text{(ICC-N)} \\
& 0 \leqslant y_n \leqslant x_n, 0 \leqslant x_1 \leqslant 1, \text{ and law of motions.}
\end{aligned}
$$

THE DESIGNER'S PROBLEM

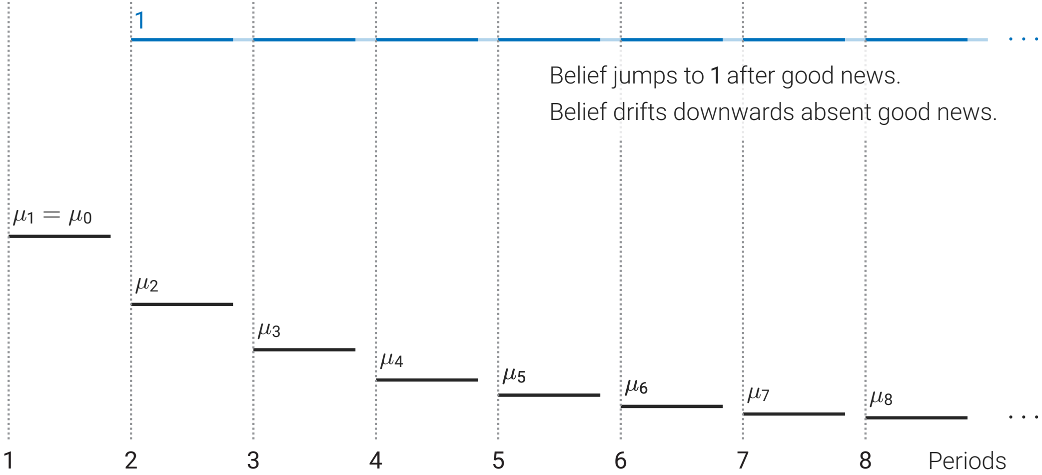
Substitute x_n and x_G out using the law of motions,

$$
\begin{aligned}
\max_{\{(x_n, y_n)\}} \quad & \sum_{n \geqslant 1} \left[\mu_n \left(1 - a + \frac{\alpha\delta}{1-\delta} \right) + a \right] y_n, \\
\text{s to.} \quad & \sum_{n \geqslant 1} \left[\mu_n \left(2 + \frac{\alpha\delta}{1-\delta} \right) - 1 \right] y_n \geqslant 0, && \text{(ICC-Y)} \\
& 2\mu_1 - 1 + \sum_{n \geqslant 1} \big(1 - \delta - [2(1-\delta) + \alpha\delta]\mu_n\big) y_n \leqslant 0, && \text{(ICC-N)} \\
& 0 \leqslant y_n \leqslant x_n, 0 \leqslant x_1 \leqslant 1.
\end{aligned}
$$

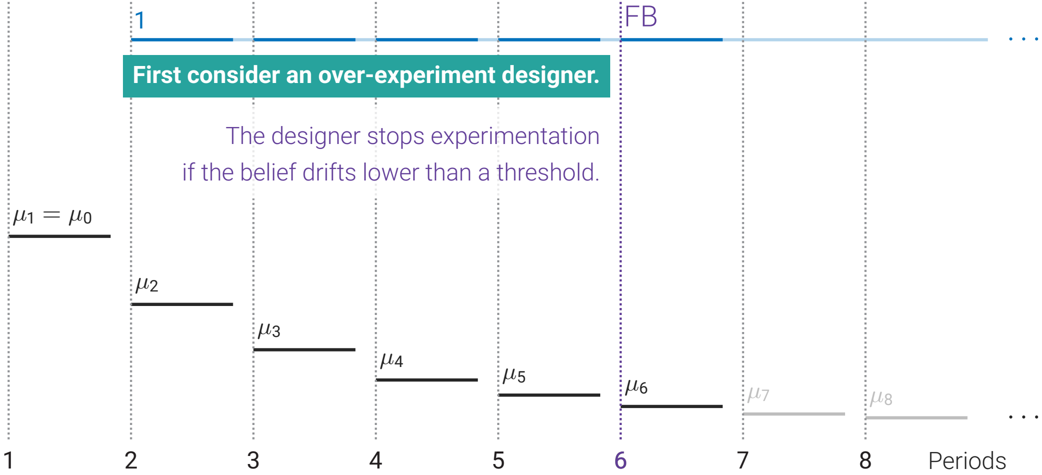
THE SECOND BEST POLICY

- The designer's second best policy contains an exploration phase (experimenting) and an exploitation phase (recommending Y if good news has arrived). Mixing, if any, only happens at the end of the exploration phase.
- When the first best is not IC, pick the correct n such that
  - $y_m = x_m$ for $m < n$, $y_m = 0$ for $m > n$.
  - $y_n \in (0, x_n]$ so that ICC-Y (if $a > -1$) or ICC-N (if $a < -1$) binds.

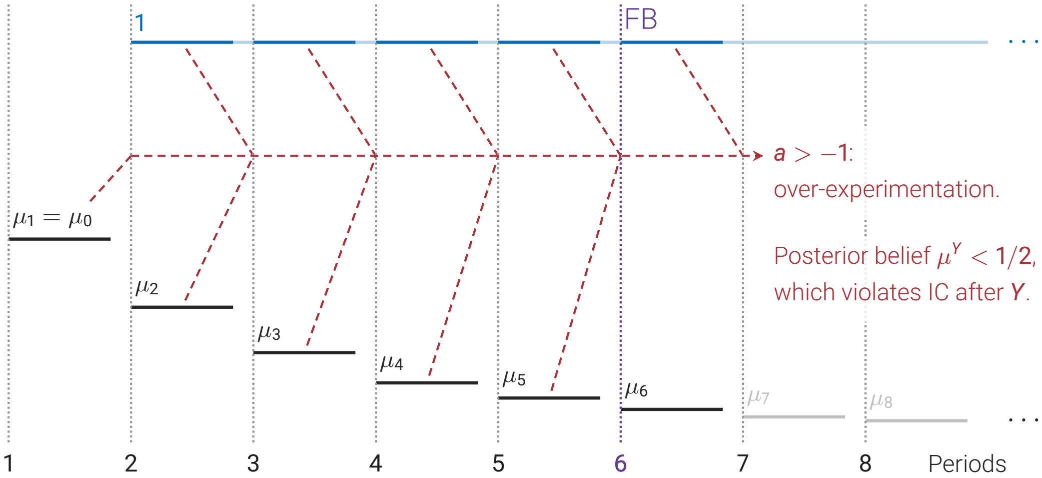
PERFECT GOOD NEWS EXAMPLE WITH $a > -1$



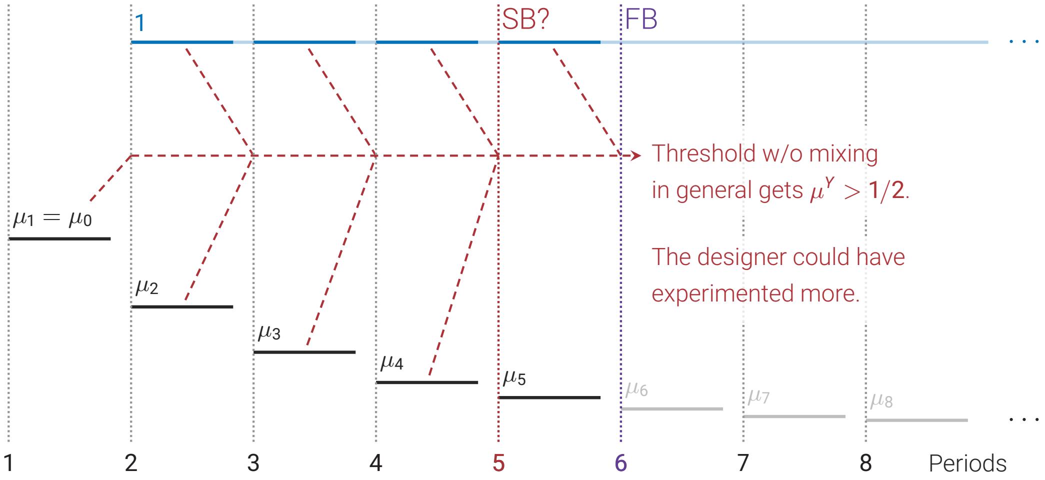
PERFECT GOOD NEWS EXAMPLE WITH $a > -1$



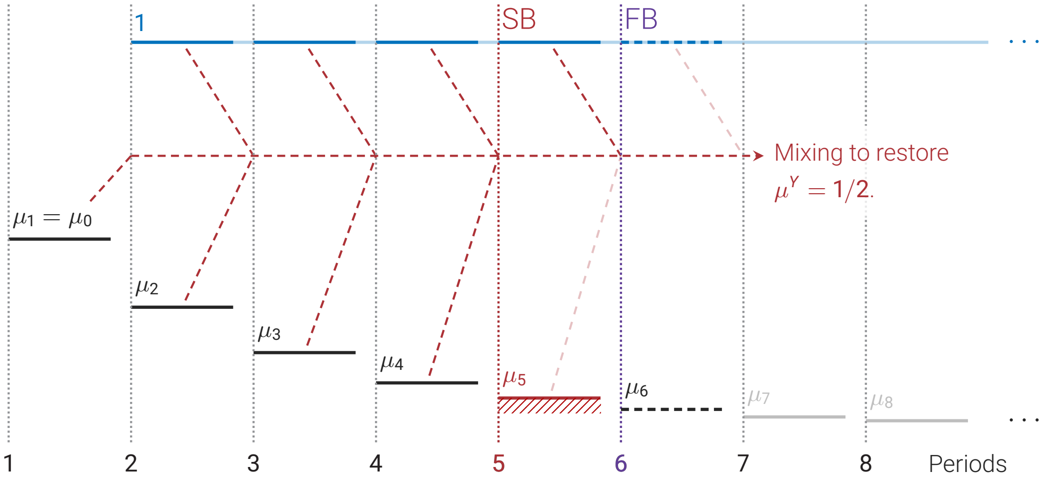
PERFECT GOOD NEWS EXAMPLE WITH $a > -1$



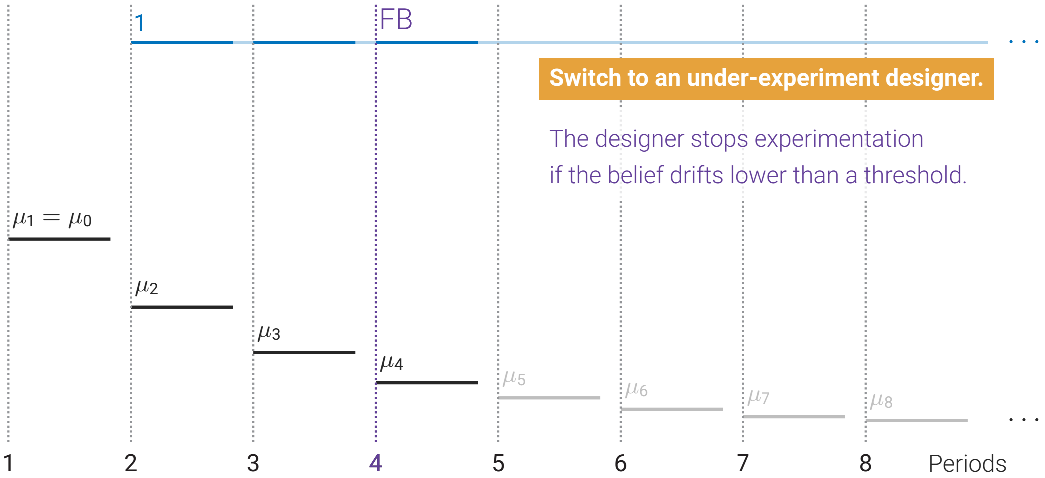
PERFECT GOOD NEWS EXAMPLE WITH $a > -1$



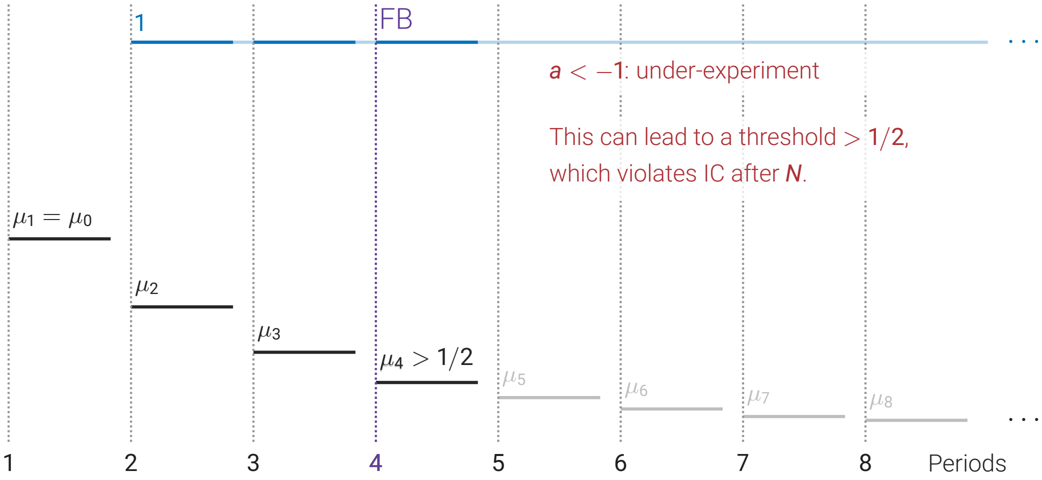
PERFECT GOOD NEWS EXAMPLE WITH $a > -1$



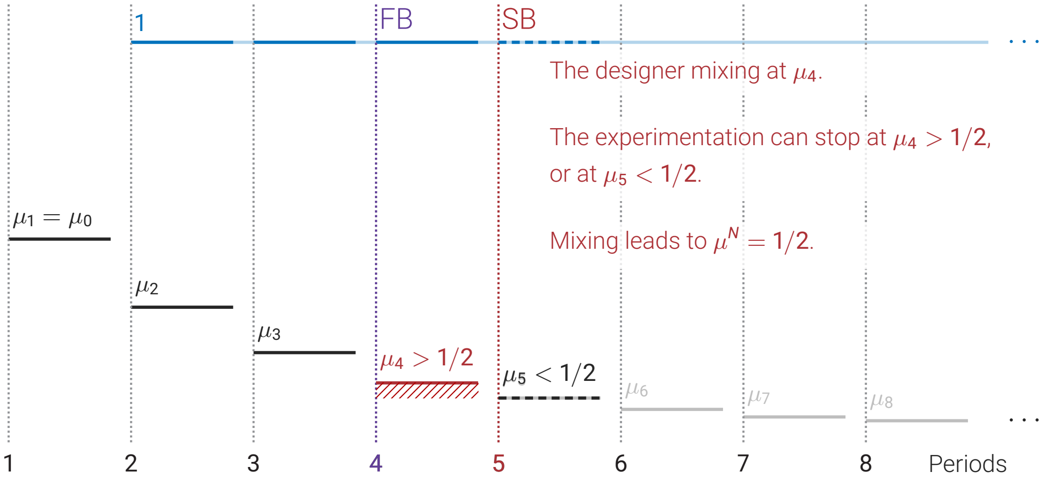
PERFECT GOOD NEWS EXAMPLE WITH $a < -1$



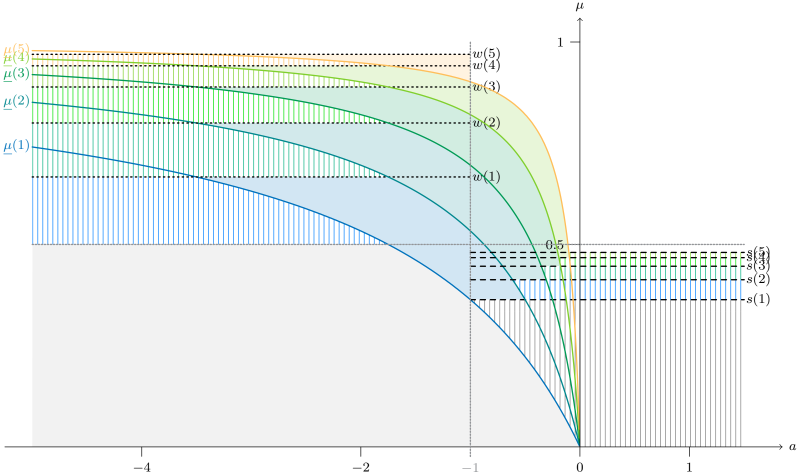
PERFECT GOOD NEWS EXAMPLE WITH $a < -1$



PERFECT GOOD NEWS EXAMPLE WITH $a < -1$



COMPARATIVE STATICS



Drawn at $\alpha = 0.5, \delta = 0.6$. Regions of more than 4 experiments are not shaded.

Vertical patterns indicate mixing in the second best policies.

Perfect Bad News

PERFECT BAD NEWS

Random bad news b under B. No news otherwise.

$$\begin{bmatrix} \mathbb{P}(b \mid G) & \mathbb{P}(n \mid G) \\ \mathbb{P}(b \mid B) & \mathbb{P}(n \mid B) \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ \beta & 1-\beta \end{bmatrix}.$$

BELIEFS AND OCCUPANCY MEASURES

- No news is good news. The belief in absence of bad news is monotonically increasing.
- Mixing, if any, happens at the beginning of the game.
- The effect of mixing is asymmetric:
  - Consumers who receive N cannot tell whether this comes from mixing at the beginning or a designer who has received bad news.
  - Consumers who receive Y realize that the mixing must have a realization Y.

The second best policies behave differently due to this asymmetry.

UNDER-EXPERIMENTING DESIGNER

Theorem.
Let $a < -1$. The designer's second best policy is either the first best policy, or a policy such that the designer mixes at μ_0 then recommends Y until bad news.

- If $\mu_0 \leqslant 1/2$, the designer can always implement the first best policy.
- If $\mu_0 > 1/2$, recommending N for sure at μ_0 is not IC. The designer instead mixes, pooling μ_0 with belief 0 arising from experimenting and bad news, so that $\mu^N = 1/2$.

OVER-EXPERIMENTING DESIGNER

Lemma.
Let $a > -1$. The designer does not mix at μ_0 in any second best policy.

- The designer wants to mix at μ_0 only if "recommending Y for sure until bad news" violates ICC-Y:

$$\mu^Y = \frac{\sum_{n \geqslant 1} \mu_n y_n}{\sum_{n \geqslant 1} y_n} < \frac{1}{2} \quad \text{with } y_1 = x_1 = 1.$$

- But by observing Y, a consumer correctly infers Y must be the realization of the mixing at μ_0, so that the mixing does not increase the posterior belief μ^Y.

OVER-EXPERIMENTING DESIGNER

Theorem. (First Best or No Experiment)
Let $-1 < a < 0$. The designer's second best policy is either the first best policy, or recommending N forever.

- The first best policy is to recommend Y for sure until bad news.
- If this is not implementable, the only policy satisfies ICC-Y is to recommend N forever.
- A product either gains popularity immediately or die at born.

PROFIT-MAXIMIZING DESIGNER

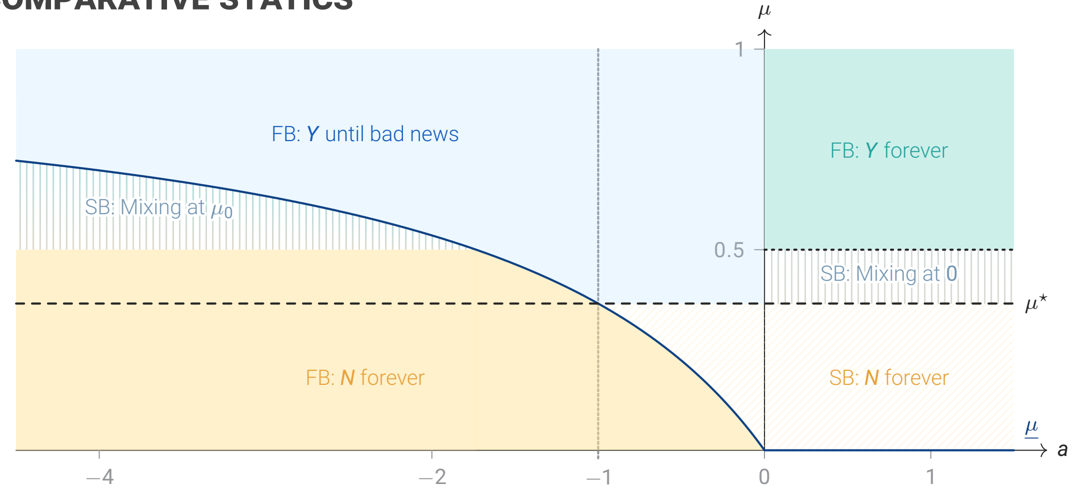
Theorem.

Let $a \geqslant 0$. The designer's second best policies take the following forms:

- If $\mu_0 \geqslant 1/2$, recommend *Y* forever.
- If $\mu^\star \leqslant \mu_0 < 1/2$, recommend *Y* until bad news, and sometimes after bad news.
- If $\mu_0 < \mu^\star$, recommend *N* forever.

- Additional middle case: For $\mu^\star < \mu_0 < 1/2$, "recommending *Y* until bad news" leaves ICC-Y slack. The designer uses this slackness to recommend known bad product to increase his payoffs.

COMPARATIVE STATICS



Drawn at $\beta = 0.6, \delta = 0.55$.

Discussions

EXTENSIONS

Consumers have noisy signals of arrival time:

- Consumers are no longer homogeneous.
- Consumers who receive signals indicating a probable early arrival are harder to be persuaded (with a harsher IC constraint).
- The designer may be able to experiment more or less depending on the consumers' realized signals of arrival time.

Consumers randomly leave no feedback:

- The designer now has a slower learning speed.
- Value of experimentation is lower, and the designer is less willing to experiment.

CONCLUSIONS

- In an environment where agents do not know how many others experimented before, an information designer can have a lot of scope to encourage or discourage experimentation.
- In particular, this is the case if designer wants to maximize expected total consumer surplus.
- For more general objective functions either first best policy implementable or one of the two IC constraints belonging to recommendations binds. Latter typically introduces random recommendation in second best.
- Second best is always a belief-threshold policy, but how it translates to the mapping between histories and recommendations depends on the learning information technology.

Appendix

Consider the case of $a > -1$ for simplicity. Suppose the first best is not IC, so that ICC-Y is binding, while ICC-N is slack.

1. Transform ICC-Y with an auxiliary designer W whose $a = -1$.
$$\max_R V(\mu_0, R) \quad \text{subject to } W(\mu_0, R) \geqslant 0.$$
2. The second best $(V, W) = (V^{SB}, 0)$ is on the Pareto frontier of the feasible (V, W)-set. Scalarize the problem using the weighting method:
$$V(\mu_0, R) + \lambda W(\mu_0, R)$$
3. Transform the problem again to a Markov problem with stage game payoff $\mu + (1 - \mu)a + \lambda(2\mu - 1)$.
4. Map the Markov threshold solution back to the original problem. Mixing at the threshold is needed to keep ICC-Y binding. [Back]