Inference from Selectively Disclosed Data

Ying Gao

November 23, 2021

Motivation

- Informed parties and decision-making bodies can be separate, with different interests
- **Hard evidence** is an important factor in generating a positive opinion, in order to win approval or buy-in
    - E.g. R&D teams and leadership, or companies and investors.
- Transparent motives to select data to support better conclusions, and selectively hide "bad" datapoints.
- Focus on problem of disclosing large datasets with many outcomes.

Background

Full disclosure: In Grossman ’81 and Milgrom ’81, senders always disclose when known to hold a single piece of evidence.

Uncertain dataset: Dye ’85 – partial disclosure with only sufficiently good news is shown.
Shin ’94, ’03 – discrete set of good and bad signals of uncertain size. Partial disclosure follows a “sanitation strategy” where only good signals are disclosed.

Dzuida ’11: disclosure with uncertain mass of good and bad signals. Presence of honest types results in “partial sanitation” equilibria under a refinement that requires outcomes are a continuous function of the measure of good signals.

Main points

Imitation is an appealing equilibrium disclosure strategy, even when data has complicated distributions.

- Senders "imitate" a desirable state by disclosing a subset of their data that mimics the distribution of data under that state.
- Separately analyze binary-state case and case of > 2 states in which senders have a choice of which states to imitate.

Outcomes of imitation equilibria with binary state are **lexicographically optimal**: maximize the payoff to senders in order of highest potential payoffs in equilibrium.

- Extends, exists and is unique with finite datasets.
- Has a foundation from stability with respect to credible announcements.

Continuous-dataset model

A sender (S) has payoffs that are increasing in a receiver's (R) beliefs about a state of the world, $\theta \in \Theta = \{1, \ldots, L\}$ with ex-ante distribution $\beta_0(\theta)$.

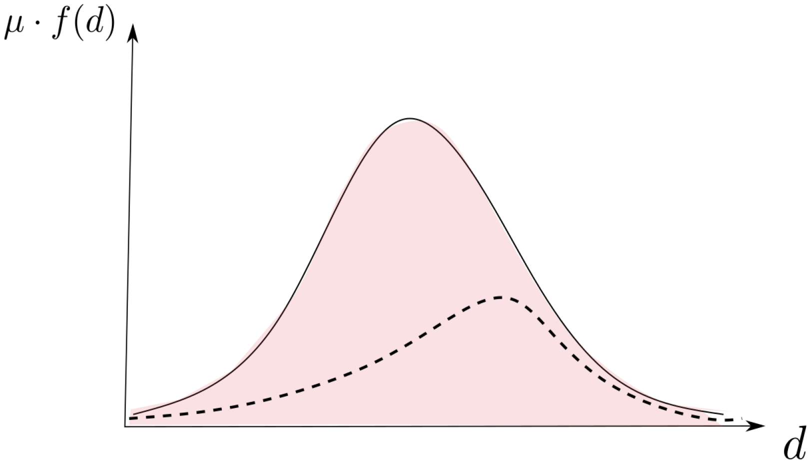
- Sender's type is (μ, θ).
- μ, distributed with continuous density $g(\mu)$ on $[0, 1]$, is the amount of data observed.
  - Assume that $g(1) = 0$.
- The distribution of the dataset has density f_θ – perfectly reveals θ.
  - $f_\theta(d)$ differs from $f_{\theta'}(d)$ on a positive-measure set for all $\theta \neq \theta'$
  - Shared support $D \subset \mathbb{R}^n$ for all θ

Receiver sees neither f_θ nor μ: relies on prior to infer both the state and sender's information endowment.

Messages

Sender can show receiver any message $(\tilde{\mu}, \tilde{f}) \subseteq (\mu, f_\theta)$.

- $(\tilde{\mu}, \tilde{f}) \subseteq (\mu, f) \Leftrightarrow \forall d, \tilde{\mu}\tilde{f}(d) \leq \mu f(d)$



Receiver and payoffs

Receiver is a Bayesian who forms a belief over types upon receiving a message, $\beta(\mu, \theta | \tilde{\mu}, \tilde{f})$. Marginal over states is $\beta(\theta | \tilde{\mu}, \tilde{f})$.

Assumption (payoff monotonicity)

S's preferences over R's marginal belief over states satisfies:

- *For all point beliefs $\mathbb{1}_\theta$ and $\mathbb{1}_{\theta'}$ where $\theta' > \theta$, $u_s(\mathbb{1}_{\theta'}) > u_s(\mathbb{1}_\theta)$.*
- *If $u_s(\beta') > u_s(\beta)$, then $u_s(\alpha\beta' + (1-\alpha)\beta)$ is continuous and increasing in α.*

Satisfied if the receiver is an expected utility maximizer who takes a 1-dimensional action $a \in \mathbb{R}$, with

- $u_R(\theta, a)$ strictly concave
- $\arg\max_a u_R(\theta, a)$ strictly increasing in θ

and the sender's payoff $u_S(a)$ is increasing in a.

Strategies and outcomes

Sender's strategy is $\sigma[\tilde{\mu}, \tilde{f}|\mu, \theta]$ and implies receiver draws inferences according to $\beta_\sigma(\theta|\tilde{\mu}, \tilde{f})$.

Restrict consideration to strategies in which:

1. $\sigma[\tilde{\mu}, \tilde{f}|\mu, \theta]$ is a measurable function of μ for every $\tilde{\mu}, \tilde{f}, \theta$
2. $\{\mu : \sigma[\tilde{\mu}, \tilde{f}|\mu, \theta] > 0\}$ is convex
3. *(For simplicity)* the support of $\sigma[\cdot|\mu, \theta]$ is finite.

An outcome of σ is

$$u_\sigma(\mu, \theta) := \sum_{\tilde{\mu}, \tilde{f} \in \text{supp}(\sigma[\cdot|\mu,\theta])} \sigma[\tilde{\mu}, \tilde{f}|\mu, \theta] u_s \left(\beta_\sigma(\cdot|\tilde{\mu}, \tilde{f})\right).$$

Equilibrium

(2), convexity of sender set participating in a message, does not restrict the set of equilibrium outcomes.

- If (μ, θ) and (μ', θ) both send the same message in equilibrium, then they and all intermediate types obtain the same payoff.
- Can rearrange their messaging strategies so that types with greater μ send "more demanding" messages; then set of μ associated with each sent message is a point or interval.

Base equilibrium concept is PBE, $(\sigma^*, \beta_{\sigma^*})$ where β_{σ^*} is Bayesian update given σ^* and σ^* is a best response to the beliefs formed by β_{σ^*}.

In general there exist many PBE. ex

Two states

Consider 2 states, $\Theta = \{H, L\}$.

- u_s is strictly increasing in $\beta(H|\cdot)$.

The sender with distribution f_θ's ability to send distribution $\tilde{f}$ is

$$r_\theta(\tilde{f}) = \max\{\tilde{\mu} : (\tilde{\mu}, \tilde{f}) \subseteq (1, f_\theta)\}.$$

In any equilibrium σ,

- $u_\sigma(\mu, \theta)$ is weakly increasing in μ
- $u_\sigma(r_L(f_H)\mu, H) \leq u_\sigma(\mu, L)$
- $u_\sigma(r_L(f_H), H) = u_s(\mathbb{1}_H)$.

Equivalent messages

Continuous-mass datasets imply some redundancy, i.e. messages that prove the same thing.

- Let $d^* \in D^* := \arg\max_d \frac{f_H(d)}{f_L(d)}$.
- For any distribution $\tilde{f}$ with $d^* \in \arg\max_d \frac{\tilde{f}(d)}{f_H(d)}$,

$$\left(\mu \frac{f_H(d^*)}{\tilde{f}(d^*)}, \tilde{f}\right) \subseteq (\mu, f_\theta) \Leftrightarrow (\mu, f_H) \subseteq (\mu, f_\theta),$$

so the types capable of sending $\left(\mu \frac{f_H(d^*)}{\tilde{f}(d^*)}, \tilde{f}\right)$ and (μ, f_H) are the same.

Equilibrium with $\tilde{f} = f_H$

Consider an imitation strategy $\bar{\sigma}$: (μ, θ) sends $(r_\theta(f_H)\mu, f_H)$.

- i.e. types (μ, L) imitate f_H by hiding data, while (μ, H) is truthful.
- This gives an equilibrium when $\frac{g(r_L(f_H)\tilde{\mu})}{g(\tilde{\mu})}$ is (weakly) monotone increasing in $\tilde{\mu}$.

If outcomes from $\bar{\sigma}$ are nonmonotone, then iron to form σ^*:

- Take $\hat{\mu} = \sup\{\tilde{\mu} : \frac{g(r_L(f_H)\tilde{\mu})}{g(\tilde{\mu})}$ decreasing in $\tilde{\mu}\}$, and find $\tilde{\mu}' < \hat{\mu} < \tilde{\mu}''$ such that

$$\frac{\beta_0 \cdot \int_{\mu'}^{\mu''} g\left(\tilde{\mu}\right) d\tilde{\mu}}{\beta_0 \cdot \int_{\mu'}^{\mu''} g\left(\tilde{\mu}\right) d\tilde{\mu} + (1-\beta_0) \int_{\mu'}^{\mu''} g\left(\frac{\tilde{\mu}}{r_L(f_H)}\right) d\tilde{\mu}} = \beta_{\bar{\sigma}}\left(H|\tilde{\mu}', f_H\right) = \beta_{\bar{\sigma}}(H|\tilde{\mu}'', f_H), \tag{1}$$

- Let (μ, θ) play $(\tilde{\mu}', f_H)$ for all $\mu \in [\frac{\tilde{\mu}'}{r_\theta(H)}, \frac{\tilde{\mu}''}{r_\theta(H)}]$, and repeat.

Equilibrium with $\tilde{f} = f_H$

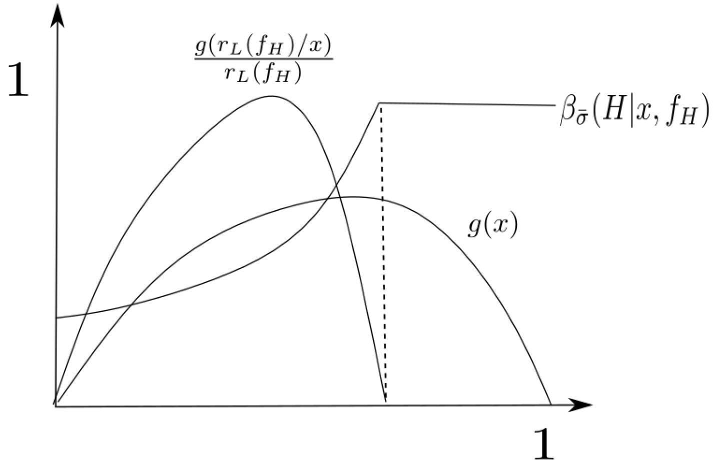


Figure: Receiver's beliefs as a function of $\tilde{m}$ when $\frac{g(r_L(f_H)\mu)}{g(\mu)}$ is monotone.

Equilibrium with $\tilde{f} = f_H$

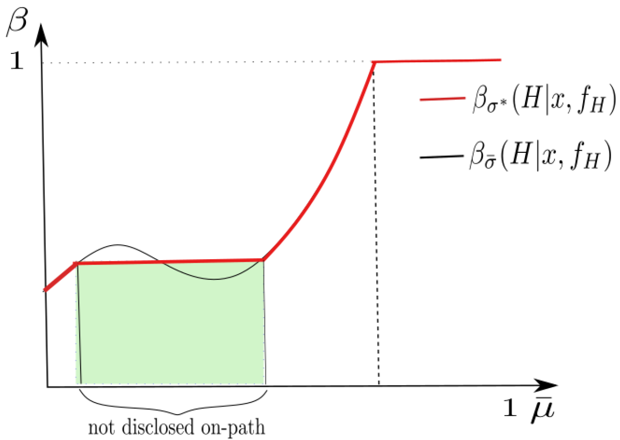


Figure: Ironed posteriors from a disclosure policy that satisfies IC for sender when $\frac{g(r_L(f_H)\mu)}{g(\mu)}$ is nonmonotone

Lexicographic optimality

An outcome $u_\sigma(\cdot)$ **lexicographically dominates** another outcome $u_{\sigma'}(\cdot)$ if there is v such that:

- $\forall u > v$, $\{(\mu', \theta') : u_{\sigma'}(\mu', \theta') \geq u\} \subseteq \{(\mu, \theta) : u_\sigma(\mu, \theta) \geq u\}$
- $\{(\mu', \theta') : u_{\sigma'}(\mu', \theta') \geq v\} \subset \{(\mu, \theta) : u_\sigma(\mu, \theta) \geq v\}$.

i.e. there exists a set of highest-payoff types in σ' that do at least as well in σ, and some do better.

Definition (lexicographic optimality)

A PBE outcome $u_\sigma(\cdot)$ is **lexicographically optimal** if it lexicographically dominates every other PBE outcome.

Lexicographic optimality

An outcome $u_\sigma(\cdot)$ is lexicographically optimal in the disclosure problem with 2 states if at each μ,

$$\begin{aligned} \sigma^* \in \arg \min_{\sigma \in \mathrm{PBE}(G)} & \left(\frac{du_\sigma(\mu r_\theta(f_H), \theta)}{d\mu} \right)^- \\ s.t.\ & \{(\mu', \theta') : u_\sigma(\mu', \theta') > u_{\sigma^*}(\mu r_\theta(f_H), \theta)\} \\ & = \{(\mu', \theta') : u_{\sigma^*}(\mu', \theta') > u_{\sigma^*}(\mu r_\theta(f_H), \theta)\} \end{aligned} \tag{2}$$

for each $\theta \in \{H, L\}$.

Lexicographic optimality

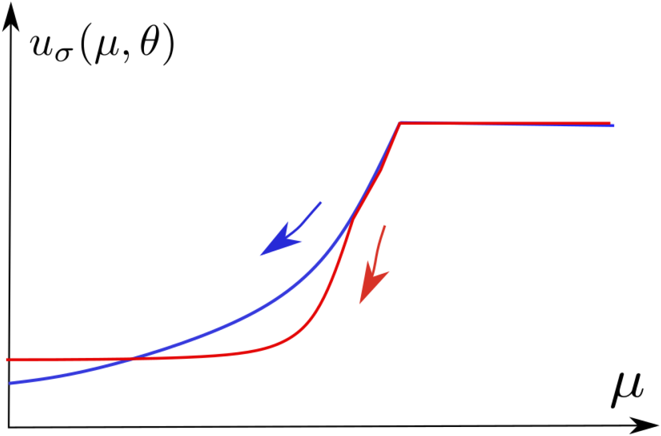


Figure: Two outcomes, as a function of μ for fixed θ.
If blue has shallower derivative at rightmost point of divergence for all θ, then blue lexicographically dominates red.

Selection of equilibrium outcome

Proposition

The imitation equilibrium outcome $u_{\sigma^*}(\cdot)$ is the unique lexicographically optimal equilibrium outcome when $|\Theta| = 2$.

- In contrast, no equilibrium attains the maximum payoff for all senders simultaneously, or is Blackwell most informative.

Also selected by a regularity condition that having more data doesn't increase propensity to send an already-feasible dataset.

Lemma

$u_{\sigma^*}(\cdot)$ is the unique outcome of equilibria in which, for all $\theta, \theta', \mu, \mu'$, and $(\tilde{\mu}, \tilde{f})$ such that $(\tilde{\mu}, \tilde{f}) \subseteq (\mu', f_{\theta'}) \subseteq (\mu, f_\theta)$,

$$\frac{\beta(\mu', f_{\theta'}|\tilde{\mu}, \tilde{f})}{\beta(\mu, f_\theta|\tilde{\mu}, \tilde{f})} \geq \frac{\pi_0(f_{\theta'})g(\mu')r_\theta(\tilde{f})}{\pi_0(f_\theta)g(\mu)r_{\theta'}(\tilde{f})}. \tag{3}$$

Finite data

Continuous-data model is stylized approximation to large datasets

- Outcomes easy to characterize, but interpretation of messaging strategies is not direct

Explicitly modeling finite data provides a robustness check, and also provides a more direct interpretation of how messages must be used in lexicographically optimal equilibria.

- Allows lexicographically optimal equilibria to be constructed without reference to the equilibrium set, rather than characterized by process of elimination.

Drawback: equilibria are hard to describe, except (possibly) in the limit.

Finite-data model

Make the model finite by letting the amount of data received be $n \sim g(n)$, with $g(n)$ supported on $\{1, \dots, N\}$.

- Also assume finite $D = \{1, \dots, k\}$ to be the domain of f_θ.

Sender's type is $t = (n_1, \dots, n_k) \in \mathcal{F}$, and the probability of each type is

$$q(t) = \frac{n!}{\Pi_{x=1}^{k} n_x!} g(n) \sum_{\theta'} \beta_0(\theta') \Pi_{x=1}^{k} f_{\theta'}(x)^{n_x}.$$

Sender does not know the state, but based on their data can evaluate the probability of state θ to be

$$\pi(\theta|t) = \frac{\beta_0(\theta) \Pi_{x=1}^{k} f_\theta(x)^{n_x}}{\sum_{\theta'} \beta_0(\theta') \Pi_{x=1}^{k} f_{\theta'}(x)^{n_x}}.$$

Partial strategies and partial updating

R's inference upon seeing $\tilde{f}$ depends only on probabilities with which senders send it, not on the rest of the strategy.

- Let $T_\sigma(\tilde{f})$ be the set of senders who play $\tilde{f}$ with positive probability
- $\hat{\sigma}_{\tilde{f}}(\tilde{f}|\cdot) : T_\sigma(\tilde{f}) \to [0,1]$ is the restriction of σ to $T_\sigma(\tilde{f})$.

$$\beta_{\hat{\sigma}}(\theta|\tilde{f}) = \frac{\sum_{t \in T_\sigma(\tilde{f})} \pi(\theta|t) q(t) \hat{\sigma}_{\tilde{f}}(t)}{\sum_{t \in T_\sigma(\tilde{f})} q(t) \hat{\sigma}_{\tilde{f}}(t)}.$$

Extend to sets of messages M, with $T_\sigma(M) = \bigcup_{\tilde{f} \in M} T_\sigma(\tilde{f})$ and $\hat{\sigma}_M : M \times T_\sigma(M) \to \mathbb{R}$.

Message classes

Given a set of types $\mathcal{T}$, define

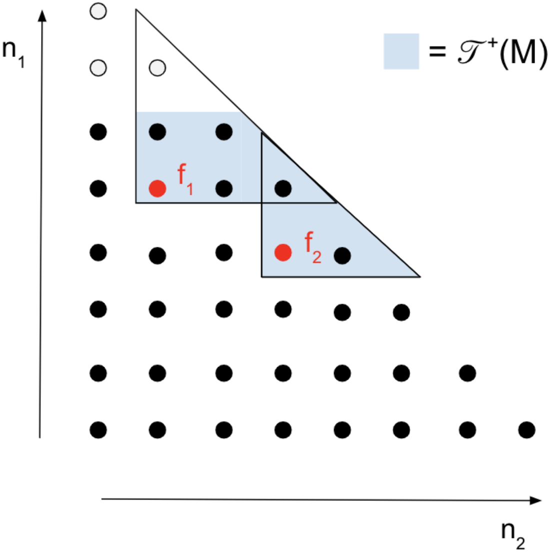
$$\mathcal{T}^+(\tilde{f}) = \{t \in \mathcal{T} : \tilde{f}_i \subset t\} \quad \text{and} \quad \mathcal{T}^+(M) = \bigcup_{\tilde{f} \in M} \mathcal{T}^+(\tilde{f})$$

as the set of types in $\mathcal{T}$ who can send $\tilde{f}$ or any $\tilde{f} \in M$, respectively.

A set of messages $M = \{\tilde{f}_1, \ldots, \tilde{f}_I\}$ is a *unifying class* in $\mathcal{T}$ if there is a partial strategy $\hat{\sigma}_M : M \times \mathcal{T}^+(M) \to \mathbb{R}$ with $\hat{\sigma}(\cdot|t) \in \Delta M$ under which eqns:

- Each type in $\mathcal{T}^+(M)$ plays messages in M with prob. 1, and M is exactly the set of messages played by types in $\mathcal{T}^+(M)$
- For any messages $\tilde{f}_i, \tilde{f}_j \in M$, payoffs are the same under $\beta_{\hat{\sigma}}$.

M and $\mathcal{T}^+(M)$



An algorithm for lexicographic optimality

Define $u(\mathcal{T})$ to be the payoff to the receiver's posterior after learning the sender's type is in $\mathcal{T}$.

Construct a lexicographically optimal strategy profile as follows.

1 Let $\mathcal{T}_1 = \mathcal{F}$, and define $\mathcal{C}_{\mathcal{T}_1}$ to be the set of unifying classes in $\mathcal{T}_1$. Take M_1 to be the union of messages in the elements of $\mathcal{C}_{\mathcal{T}_1}$ that yield the highest payoff to participating senders:

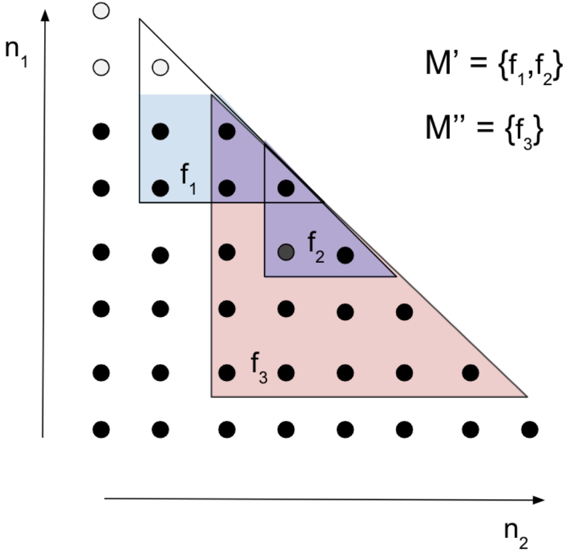
$$M_1 = \bigcup \{\arg \max_{M \in \mathcal{C}_{\mathcal{T}_1}} u(\mathcal{T}_1^+(M))\}.$$

Can show that M_1 is itself a payoff-maximizing unifying class of messages in $\mathcal{T}_1$, and thus the largest one.

Lemma

$M_1 \in \arg\max_{M \in \mathcal{C}_{\mathcal{T}_1}} u(\mathcal{T}_1^+(M))$.

Unique largest payoff-maximizing unifying class



An algorithm for lexicographic optimality

2 For $m = 2$ onwards, restrict the set of types to $\mathcal{T}_m = \mathcal{T}_{m-1} \setminus \mathcal{T}^+_{m-1}(M_{m-1})$, and create the class

$$M_m = \bigcup \{\arg \max_{M \in \mathcal{C}_{\mathcal{T}_m}} u(\mathcal{T}^+_m(M))\}.$$

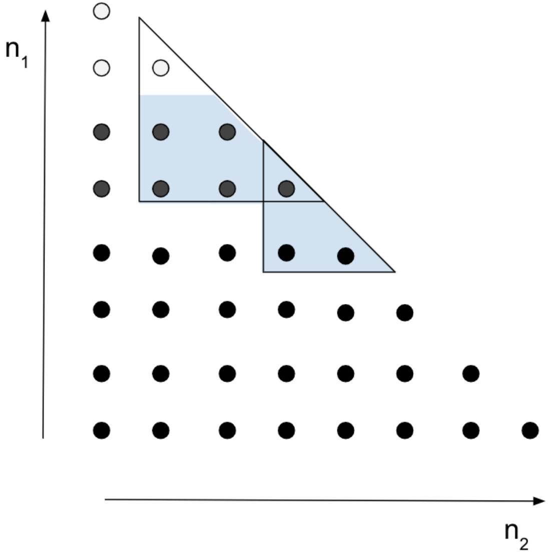
3 Continue until $\mathcal{T}_m \setminus \mathcal{T}^+_m(M_m) = \emptyset$, and define σ^* by $\sigma^*(\tilde{f}|t) = \hat{\sigma}_{M_m}(\tilde{f})$ where M_m is the class containing $\tilde{f}$.

Proposition

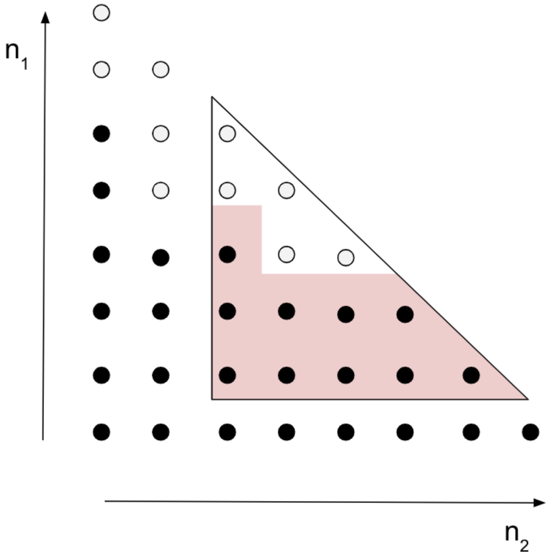
σ^* is an equilibrium.

This requires that $u(\mathcal{T}^+_m(M_m)) \geq u(\mathcal{T}^+_{m+1}(M_{m+1}))$ for all m.

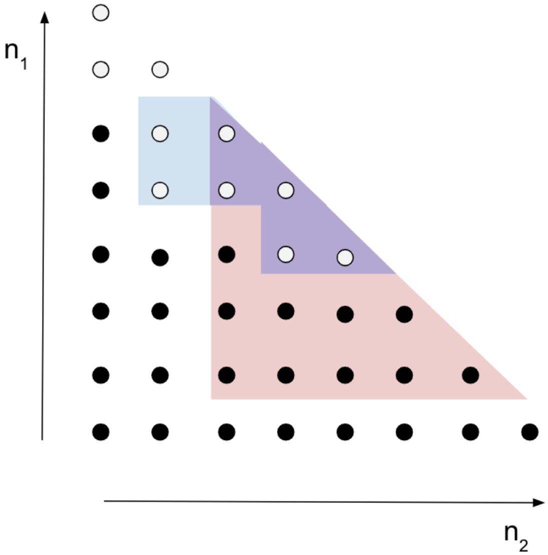
Equilibrium



Equilibrium



Equilibrium



Inclusive announcements

Definition

Given an outcome u_σ, a set of types T has a **credible inclusive announcement** to play message set M for payoff v if

- There is partial strategy $\hat{\sigma}_M : M \times T \to \mathbb{R}$ such that $\forall \tilde{f} \in M$, $\sum_i \hat{\sigma}_{\tilde{f} \in M}(\tilde{f}|t) = 1$ for $t \in T$ and $u_s(\beta_{\hat{\sigma}_M}(\cdot|\tilde{f})) = v$.
- $T = \{t : u_\sigma(t) \leq v \text{ and } \exists \tilde{f} \in M \text{ s.t. } \tilde{f} \subseteq t\}$.
- There is some $t \in T$ with $u_\sigma(t) < v$.

- May involve reinterpreting messages already in play in σ.
- Related to credible announcements (Matthews et al '91), but differs in requiring participation of *all* weakly better-off types.

Claim

An outcome u_σ is lexicographically optimal if and only if no set of types has a credible inclusive announcement under it.

Properties of u_{σ^*}

Proposition

u_{σ^*} is the lexicographically optimal equilibrium outcome.

Hard to solve for u_{σ^*} in general, but feasible when restricting to 2 states, $\Theta = \{L, H\}$ and 2 outcomes, $X = \{l, h\}$, with $p(h|H) > p(h|L)$.

- Optimal strategy is to hide all ls, and disclose a subset of hs.
- Sps. a sequence of finite data-generating functions limits to the continuous dataset-generating function $g(\mu)$ as $N \to \infty$.

  Then the limit of outcomes $u_{\sigma_N^*}$ under these models limits to the imitation equilibrium outcome under continuous datasets.

Conjecture: With > 2 signal realizations, u_{σ^*} also limits to the imitation equilibrium outcome.

> 2 states

Return to continuously-distributed datasets to characterize outcomes when $|\Theta| = J > 2$.

- In an imitation equilibrium, on-path strategies are

$$\{(\tilde{\mu}, f_\theta)\}_{\mu \in [\underline{\mu}, \bar{\mu}], \theta \in \Theta}.$$

- $u_s(\beta(\cdot|\tilde{\mu}, f_\theta))$ is (weakly) increasing in $\tilde{\mu}$ for each $\theta \in \{1, \ldots, J\}$.*
- Optimization by sender implies (μ, θ) should choose to imitate state

$$\tilde{\theta} \in \arg\max_{\tilde{\theta}'} u_s(\beta_\sigma(\cdot|r_\theta(f_{\tilde{\theta}'})\mu, f_\theta))$$

*At least on path; off-path beliefs can always be made to satisfy weak monotonicity in $\tilde{\mu}$.

Burden of proof

- Can summarize with a vector-valued burden-of-proof function, $\hat{\mu}(u) = (\hat{\mu}_1(u), \ldots, \hat{\mu}_J(u))$.
  - $\hat{\mu}_j(u)$ is the amount of data distributed f_j necessary to achieve payoff u.
  - Each sender need only meet a component of $\hat{\mu}(u)$ in order to obtain u, so

$$u_\sigma(\mu, \theta) = \max\{u : \exists j \text{ s.t. } (\hat{\mu}_j(u), j) \subseteq (\mu, \theta)\}$$

- $\hat{\mu}_j(u)$ has domain $[u_s(\mathbb{1}_1), u_s(\mathbb{1}_j)]$ and inverse $u_j(\tilde{\mu})$ that gives the payoff to message $(\tilde{\mu}, f_j)$.

Burden of proof

Theorem

$\exists$ a unique vector-valued function $\hat{\mu}(u) : [0, u_s(\mathbb{1}_{\theta=1})] \to \mathbb{R}^J$ s.t.

1. *$u_j(\tilde{\mu})$ is continuous and (weakly) increasing in $\tilde{\mu}$ for all j.*
2. *There is a strategy σ^{Im} with $\sigma^{Im}(\mu, f_\theta)$ supported on*

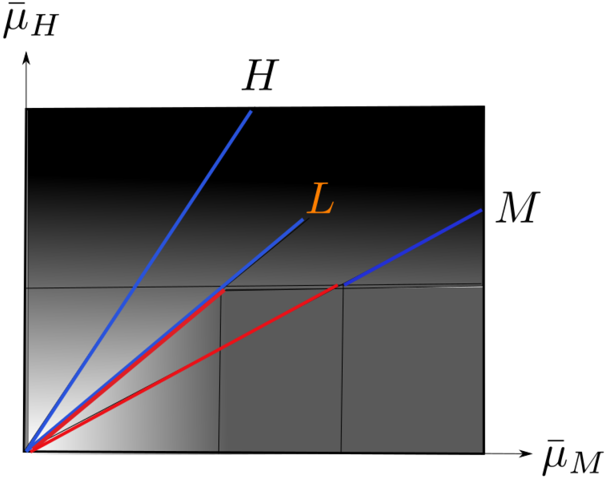
$$\tilde{S}_\theta(\mu) = \{(\hat{\mu}_j(u), f_j) : j \in \arg\max_j u_j(\mu r_\theta(f_j))\}$$

with $\sigma^{Im}[\hat{\mu}_j(u), f_j | \hat{\mu}_j(u), j] = 1$ for all j such that $u_s(\mathbb{1}_j) \geq u$ and such that for each u and j,

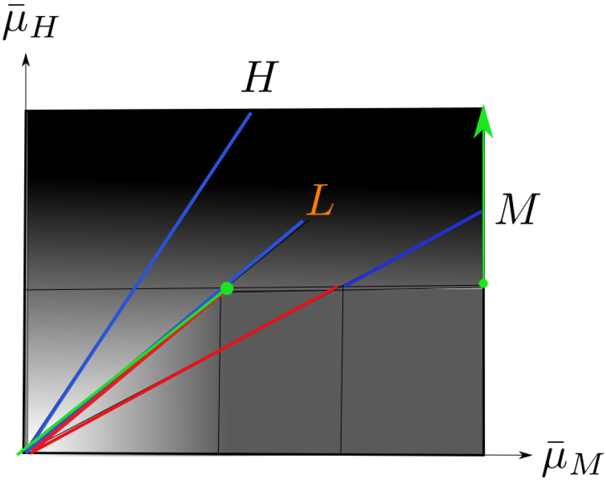
$$u_{\sigma^{Im}}(\hat{\mu}_k(u), f_k) = u.$$

σ^{Im} is an equilibrium sender strategy profile, and $\hat{\mu}(\cdot)$ is the corresponding burden-of-proof function.

Equilibrium



Equilibrium



Comments

- $\hat{\mu}$ and σ^{Im} can be constructed iteratively. A summary:
  - Payoffs $u \in [u_s(\mathbb{1}_{j-1}), u_s(\mathbb{1}_j)]$ require imitating a state in $\{j, \ldots, J\}$.
  - Initial condition $\hat{\mu}_J(\mathbb{1}_J) = \max_{j<J}\{r_j(f_J)\}$.
  - Rate of change of $u_k(\tilde{\mu})$ pinned by change in ratios of $g(\frac{\tilde{\mu}}{r_\theta(f_k)})$; rate of change of $\sigma^{Im}(r_\theta(f_j)\mu, f_j|\mu, \theta)$ in μ for mixing types is pinned by indifference.
- If (μ, θ) hides data under σ^{Im}, then it obtains payoff strictly greater than $u_s(\mathbb{1}_\theta)$; if not, then it does not.

Conjecture: $u_{\sigma^{Im}}$ is lexicographically optimal.
Conjecture 2: $u_{\sigma^{Im}}$ is the limit of finite-data equilibria with > 2 states and data mass distributions limiting to $g(\mu)$.

PBE (binary state)

PBE is $(\sigma^*, \beta_{\sigma^*})$ where

- supp $\sigma^*[\cdot|\mu,\theta] \subseteq \arg\max_{\tilde{\mu},\tilde{f}} u_s(\beta_{\sigma^*}(\cdot|\tilde{\mu},\tilde{f}))$ s.t. $(\tilde{\mu},\tilde{f}) \subset (\mu, f_\theta)$.
- $\beta_{\sigma^*}(\theta|\tilde{\mu},\tilde{f}) = \frac{\beta_0(\theta)\int_\mu g(\mu)\sigma^*(\tilde{\mu},\tilde{f}|\mu,\theta)d\mu}{\sum_{\theta'\in\Theta}\beta_0(\theta')\int_\mu g(\mu)\sigma^*(\tilde{\mu},\tilde{f}|\mu,\theta')d\mu}$

PBE with bad off-path beliefs:

- There exists a PBE in which types $(\mu \geq r_L(f_H), H)$ are honest, while all types under state L and types $(\mu < r_L(f_H), H)$ all disclose nothing.

PBE with bad on-path beliefs:

- If $\beta(\theta|\tilde{\mu},\tilde{H})$ is the same for all $\tilde{\mu} \in [\tilde{\mu}', \tilde{\mu}'']$, then some senders under state H that are capable of sending e.g. $(\tilde{\mu}'', H)$ may instead send $(\tilde{\mu}, H)$ for some $\tilde{\mu} < \tilde{\mu}''$.
- Some mixed strategy of this form in turn supports a uniform belief on $[\tilde{\mu}', \tilde{\mu}'']$.

Message classes

A set of messages $M = \{\tilde{f}_1, \ldots, \tilde{f}_I\}$ is a *unifying class* in $\mathcal{T}$ if there is a partial strategy $\hat{\sigma}_M : M \times \mathcal{T}^+(M) \to \mathbb{R}$ with $\hat{\sigma}(\cdot|t) \in \Delta M$ under which:

A. $\sum_i \hat{\sigma}_M(\tilde{f}_i|t) = 1$ for all $t \in \mathcal{T}^+(M)$.

B. For each $\tilde{f}_i$, there is $T_{\hat{\sigma}_M}(\tilde{f}_i)$ such that $\hat{\sigma}_M(\tilde{f}_i|t) = 0$ for $t \notin T_{\hat{\sigma}_M}(\tilde{f}_i)$, and $\mathcal{T}^+(M) = \bigcup_i T_{\hat{\sigma}_M}(\tilde{f}_i)$.

C. $u_s(\beta_{\hat{\sigma}_M}(\cdot|\tilde{f}_i)) = u_s(\beta_{\hat{\sigma}_M}(\cdot|\tilde{f}_j))$ for all i, j.

Credible announcements

A *weakly credible announcement* is $\langle(\tilde{\mu}, \tilde{f}), (\tau, T)\rangle$ such that:

1. For all types (μ, θ) in T, and all messages $(\tilde{\mu}, \tilde{f})$ in the support of $\tau(\cdot|\mu, \theta)$,
    1. $u_s[(\tilde{\mu}, \tilde{f}), (\tau, T)|\mu, \theta] \geq u_s[\sigma|\mu, \theta]$, with strict inequality for some $(\mu, \theta) \in D$.
    2. $u_s[(\tilde{\mu}, \tilde{f}), (\tau, T)|\mu, \theta] \geq u_s[(\tilde{\mu}', \tilde{f}'), (\tau, T)|\mu, \theta]$ for any $\tilde{\mu}', \tilde{f}'$ played with positive probability under τ.
2. For all types (μ, θ) not in T, and all messages $(\tilde{\mu}, \tilde{f})$ announced with positive probability under τ, $u_s[(\tilde{\mu}, \tilde{f}), (\tau, T)|\mu, \theta] \leq u_s[\sigma|\mu, \theta]$.