Buying Opinions

Mark Whitmeyer (Joint with Kun Zhang)

SEA 22

November 20, 2022

- Many situations in which decision makers pay for advice. Two particularly relevant:
  - Scouts/Headhunters.
  - Consulting Firms.
- A bilateral contracting scenario: principal (P) pays for an agent's (A's) advice.
- Key features:
  - A's information acquisition is flexible, costly and private.
  - A's findings are unverifiable: A acquires information before sending a cheap-talk message to P.
  - P can condition contract on A's message and state.
  - At any point (*ex ante* and interim) A can decline to participate/report and take outside option.
- Standard MH decomposition:
  1. How to efficiently implement a posterior distribution.
  2. What distribution to implement.

Preview of Findings

- P can implement any distribution over posteriors.
- An agent's optimal learning pins down the relative incentives (our version of IC): P's optimization problem simplifies to n-variable problem (n the number of states).
- When A is risk neutral and negative transfers allowed (no l.l.), P can implement any distribution at first-best cost. Overall problem just like the single-agent problem. **Selling the project to the agent does not work!**
- Characterization of optimal implementation:
  - l.l. and risk-neutral A: first-best implementation for sufficiently contracted distributions. Rents for A if first-best infeasible.
  - No l.l. and risk-averse A: first-best infeasible. Rents for A (generically).

Related Work

- Rappoport & Somma (2017): posteriors are contractible.
  - Hard (them) versus soft (us) information.
- Yoder (2022): posteriors are contractible, agent's marginal cost of information (κ) is private information.
  - Screening is now important;
  - Contracting on experiment versus posteriors.
- Zermeño (2011), Clark & Reggiani (2021): decision-making delegated to the agent;
  - Can payoffs depend on true state?
  - Decomposition of Pareto optimal contracts.

Model

- P has decision problem in which she chooses $a \in \mathcal{A}$, where $\mathcal{A}$ is compact.
- P's utility is $u(a, \theta)$, where $\theta \in \Theta$ is the unknown state, $|\Theta| = n$. u is continuous in a.
- P and A share common (WLOG, full support) prior $\mu \in \Delta(\Theta)$.
- A can acquire information, flexibly, subject to a cost:
  - A chooses any Bayes-plausible $F \in \Delta\Delta(\Theta)$ and incurs $C(F) = \kappa \int_{\Delta(\Theta)} c dF$
  - $\kappa > 0$ scales the cost
  - $c : \Delta(\Theta) \to \mathbb{R}_+$ is strictly convex, 2x differentiable, bounded on $\text{int}\,\Delta(\Theta)$, and $c(\mu) = 0$.
  - Class includes entropy (Sims 2003), log-likelihood (Pomatto, Strack & Tamuz 2020), and quadratic (Tsallis 1988).

Model

- After acquiring information, A sends a message to P, who then takes an action.
- True state is *ex post* observable and contractible.
- Contract is a pair (M, t):
  - A compact set of messages M available to the agent, and
  - A transfer $t: M \times \Theta \to \mathbb{R}$ ($t: M \times \Theta \to \mathbb{R}_+$ if limited liability).
- P's payoff is quasi-linear in the transfer.
- A's payoff is additively separable in her utility from transfer and cost of acquiring information.
- A values transfer according to $v(\cdot)$, which is continuously differentiable, strictly increasing, weakly concave, and satisfies $v(0) = 0$.
- A has outside option $v_0 \geq 0$ (P gets disutility $> v_0$ if A takes o.o.).
  - A can take this after (M, t) is proposed or after acquiring information.

Timing

1. P proposes (M, t);
2. If A doesn't accept, the game ends; otherwise the agent chooses F;
3. A posterior $\mathrm{x} \in \Delta(\Theta)$ is drawn from F, which is privately observed by A;
4. A chooses whether to report, and if she reports, she sends a message $m \in M$;
5. P takes an action $a \in A$;
6. The true state $\theta \in \Theta$ realizes;
7. P gets $u(a, \theta) - v^{-1}(t(m, \theta))$, and A gets $t(m, \theta) - c(F)$.

First-Best Benchmark

- Write the P (expected) gross payoff as a function of the posterior x, $V(x)$.
- $V(\cdot)$ is convex (and if A is finite it is piecewise affine and convex).
- Denote the set of Bayes-plausible distributions over posteriors by $\mathcal{F}(\mu)$.
- It is a convex and compact subset of $\Delta\Delta(\Theta)$.
- If the principal controlled the information acquisition herself, she would solve

$$\max_{F \in \mathcal{F}(\mu)} \int (V - \kappa c)\, dF \, .$$

- First-best: P can observe A's choice of F and specify transfer $t : \Delta\Delta(\Theta) \to \mathbb{R}_+$. Cost of acquiring information is $v^{-1}(v_0 + C(F))$.

Inducing a Distribution

- WLOG for any distribution P wants to implement, M is the support of the distribution.
- Following Caplin, Dean, & Leahy (2022), decision problem (μ, D, w) as the choice over a compact set of actions D given the prior μ over states in Θ, and $w \colon D \times \Theta \to \mathbb{R}$ is the decision maker's utility function.
- Given a decision problem (μ, D), the DM chooses a Bayes-plausible distribution over posteriors G and an action strategy $\sigma \colon \text{supp}(G) \to \Delta(D)$.
- A contract (M, t) induces a decision problem (μ, M, t) of the agent.
- A distribution, F, is implementable if there exists a contract (M, t) such that $M = \text{supp}(F)$, and the agent's optimal strategy is $\left(F, \{\delta_x\}_{x \in \text{supp}(F)}\right)$.

The Agent's Decision Problem

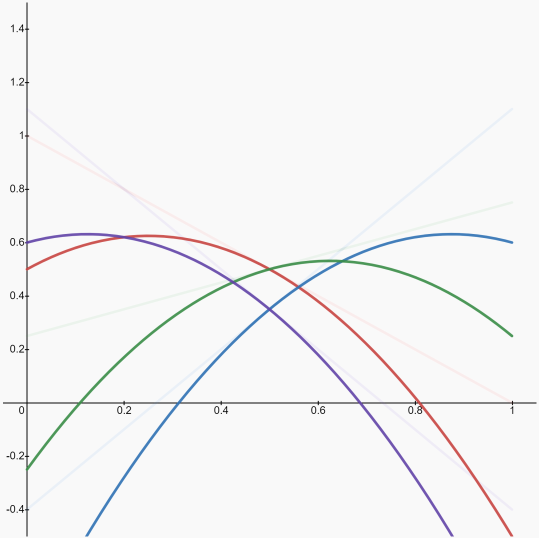
- For any $m \in M$, define A's *net utility* $N(\mathbf{x} \mid m)$:

$$N(\mathbf{x} \mid m) = \sum_{i=1}^{n-1} x_i t(m, \theta_i) + \left(1 - \sum_{i=1}^{n-1} x_i\right) t(m, \theta_n) - \kappa c(\mathbf{x}) \; ,$$

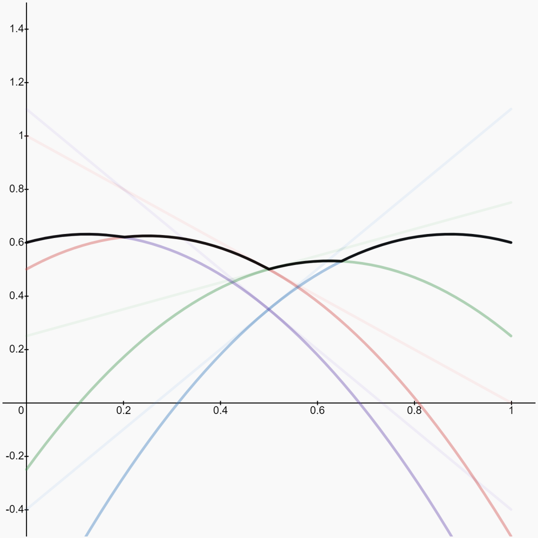
where x_i is the i-th entry of $\mathbf{x}$.

- A chooses a distribution over posteriors G to maximize her value function $W(\mathbf{x}) = \max_{m \in M} N(\mathbf{x} \mid m)$.
- A's optimal G is given by concavifying W: affine function $f_{\mathcal{H}}(\mathbf{x}) : \Delta(\Theta) \to \mathbb{R}$ intersects W at support of G $(= F)$.
- Set of intersection points of $f_{\mathcal{H}}$ and W is $P_{(M,t)} \Rightarrow F$ can be implemented by (M, t) only if $\operatorname{supp} F = P_{(M,t)}$.

Two-State Illustration



Two-State Illustration



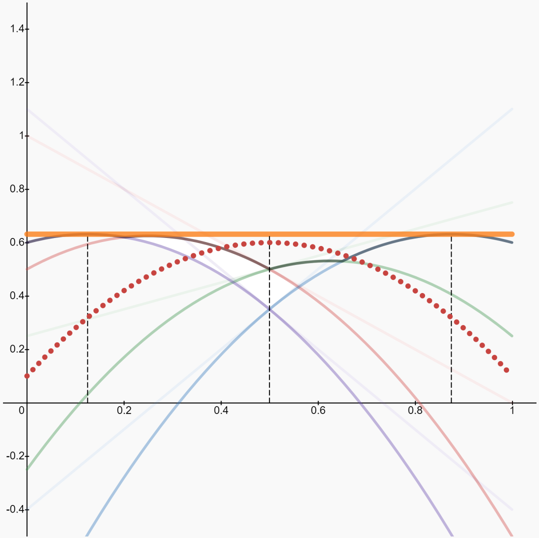
Two-State Illustration

But What About the Outside Option!?!?!?!

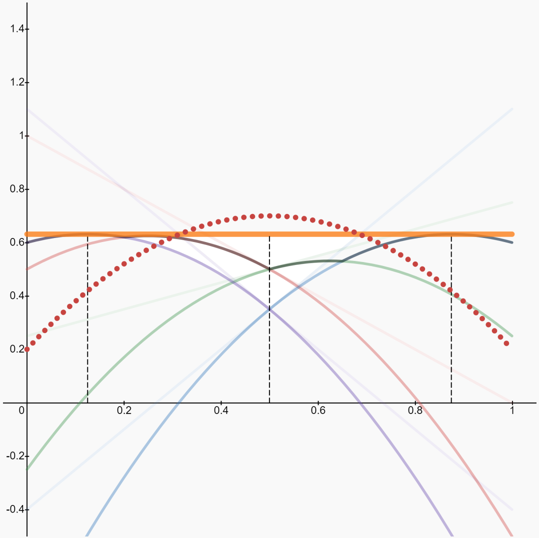
- The condition above needn't be sufficient for implementation.
- The contract must also prevent A from walking away *at any point* in the interaction.
- No double deviations (learn differently and walk away at some belief):

$$f_{\mathcal{H}}(\mathbf{x}) \geq v_0 - \kappa c(\mathbf{x}) \quad \text{for all} \quad \mathbf{x} \in \Delta(\Theta)\,. \tag{IR}$$

Ok



Not Ok



How to Implement a Distribution?

Lemma 1 A contract (M, t) implements a distribution F if and only if

1. **IC:** $\text{supp}(F) = P_{(M,t)}$; and
2. **IR:** Constraint *IR* holds; and
3. **LL:** If imposed, $t(m, \theta) \geq 0$ for all $\theta \in \Theta$ and $m \in M$.

- Without interim participation, *IR* is just $f_{\mathcal{H}}(\mu) \geq v_0$.
- Salvage value? Replace v_0 with usc curve $\rho(x)$.

Two Preliminary Results

Lemma 2 If F is a distribution over posteriors with $|\text{supp}(F)| \leq n$ and $\text{supp}(F) \subsetneq \text{int}\,\Delta(\Theta)$, there exists a contract (M, t) that implements F, and the expected cost to the principal is finite.

Corollary 3

1. Every $F \in \mathcal{F}(\mu)$ with $\text{supp}(F) \subsetneq \text{int}\,\Delta(\Theta)$ can be induced at a finite cost.
2. WLOG, P only induces distributions with support on at most n points.

A Big Simplification

- For each state $k = 1, \ldots, n$, define $\Omega^k(i,j) := t_i^k - t_j^k$ $(i,j = 1, \ldots, s)$.
- Each $\Omega^k(i,j)$ specifies the difference between the payoff to the agent from sending any (on path) message i versus message j in state k.

Theorem[Identification/Non-identification] Given a distribution over posteriors F chosen by an agent and an information acquisition cost function c, only the relative incentives $\left(\Omega^k(i,j)\right)_{i,j=1,\ldots,s;k=1,\ldots,n}$ are identified.

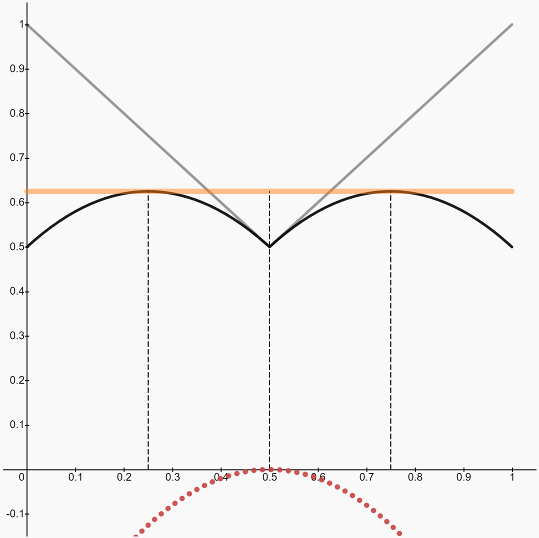
- For each state k, P fixes benchmark message $j(k)$, then chooses $\left(t_{j(k)}^k\right)_{k=1}^n$; the payoff to A from sending message $j(k)$ in state k

No Limited Liability

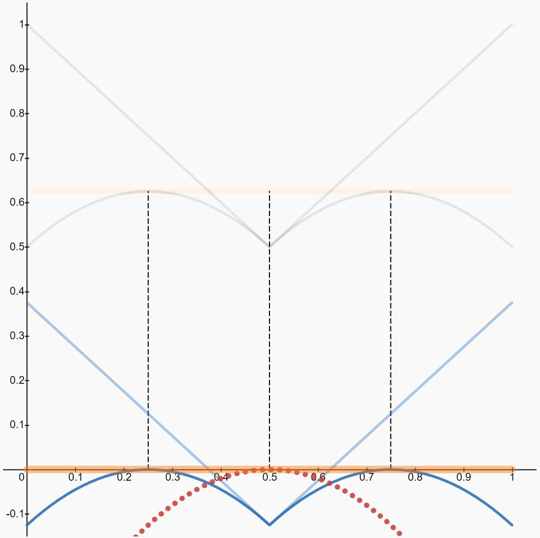
Risk-Neutral A and No L.L.

- Efficient (first-best) implementation requires $f_{\mathcal{H}}(\mu) = v_0$.
- Thus, Constraint IR must bind at $\mathbf{x} = \mu$.
- **Selling the project to the agent?**

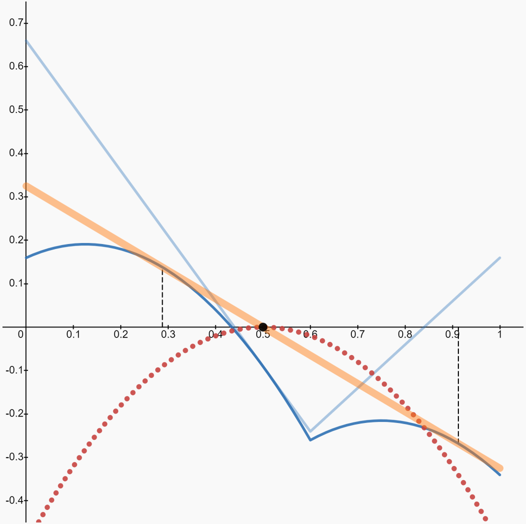
STP2TA Step 1



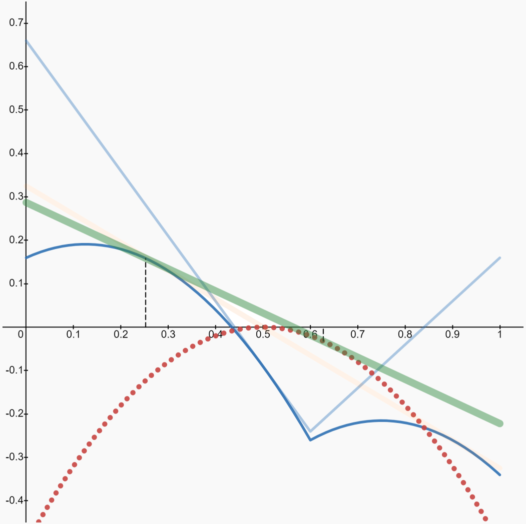
STP2TA Step 2



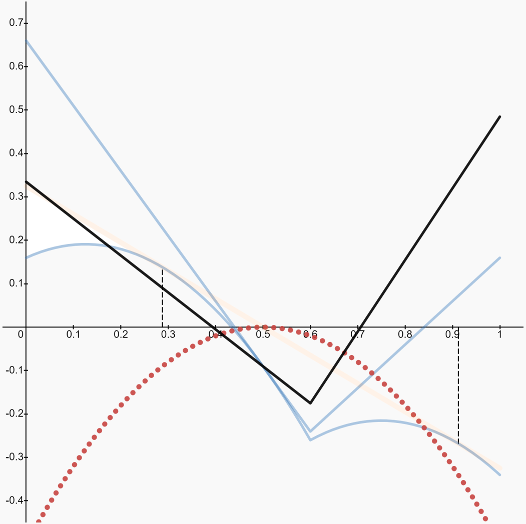
Perturbing P's Payoff



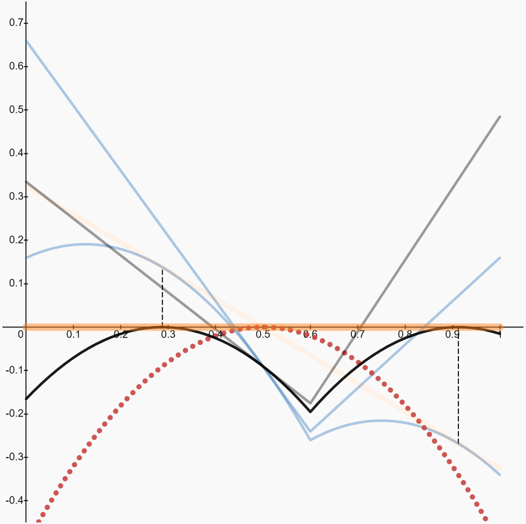
Disaster!



A Different Contract



Success!



Risk-Neutral A and No L.L. = Success

- No interim IR $\Rightarrow$ selling the project works. Key thing: $f_{\mathcal{H}}(\mu) = v_0$.
- Interim IR $\Rightarrow$ selling the project doesn't work generically. Now need $f_{\mathcal{H}}$ *tangent* to $v_0 - \kappa c$ at μ.

Proposition 5 If A is risk neutral and not protected by l.l., every (feasible) F with $\text{supp}(F) \subseteq \text{int}\,\Delta(\Theta)$ can be implemented efficiently.

- Not a *shoot the agent contract*. Penalties may be mild.
- If v_0 is sufficiently large (or implemented distribution sufficiently low in Blackwell order), l.l. satisified.

Three Notes

- Analog not true in classical setting. There interim IR = l.l. $\Rightarrow$ rents for A.
- Connection to dynamic information acquisition: extra dimension (time) not used/needed!
- Result holds even if set of feasible distributions is restricted (some subset of $\mathcal{F}_\mu$).

Limited Liability

First Observation for Low Outside Option

- Unless the implemented distribution is δ_μ, A must get rents.
- Intuition: just think of A's payoff gross of info costs

$$\max_{m \in M} \left\{ \sum_{i=1}^{n-1} x_i t(m, \theta_i) + \left(1 - \sum_{i=1}^{n-1} x_i\right) t(m, \theta_n) \right\} > v_0 ,$$

for all v_0 sufficiently close to 0.

- A's net payoff (gross minus $-\kappa c$) must therefore lie strictly above $v_0 - \kappa c$.

Proposition 6 For each state $k = 1, \ldots, n$, there exists $j^*(k)$ such that $t\left(\mathrm{x}_{j^*(k)}, \theta_k\right) = 0$, and all other transfers are determined by optimal learning.

Limited Liability & Risk-Neutral A (2 States)

Proposition 7 P can implement $\{x_L, x_H\}$ efficiently if and only if $v_0/\kappa \geq \eta(x_L, x_H)$.

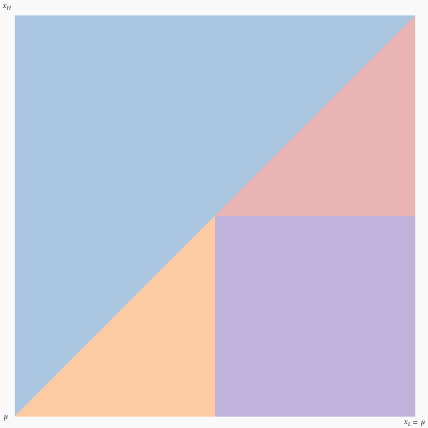
- Function η is (smoothly) decreasing in x_L and increasing in x_H. Equals 0 when $x_L = x_H = \mu$ (degenerate distribution).
- LHS increasing in o.o., decreasing in cost of information $\Rightarrow$ easier to implement first-best when o.o. is high or information is cheap.

Full Characterization (2 States)

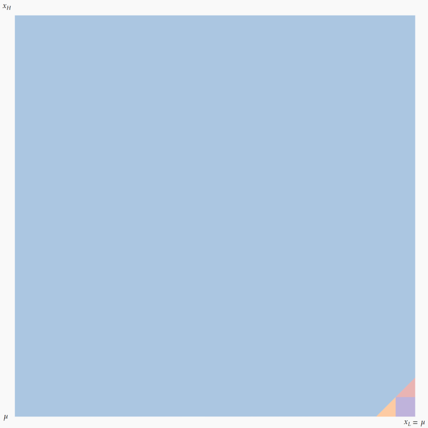
Proposition 8

1. $\{x_L, x_H\}$ can be implemented efficiently (and Constraint *IR* binds); or
2. $\{x_L, x_H\}$ cannot be implemented efficiently; and either
   - 2.1 Constraint *IR* binds and $t(x_L, \theta_1) = 0$; or
   - 2.2 Constraint *IR* binds and $t(x_H, \theta_0) = 0$; or
   - 2.3 Constraint *IR* does not bind and $t(x_L, \theta_1) = t(x_H, \theta_0) = 0$.

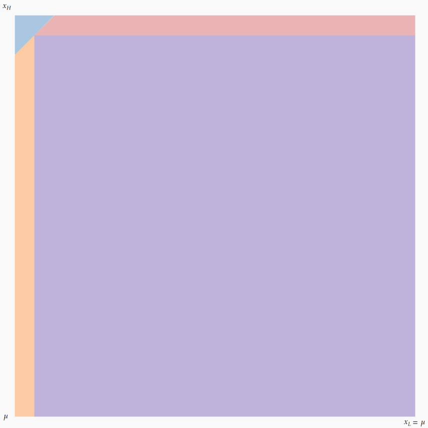
Entropy Reduction Cost: Moderate O.O.



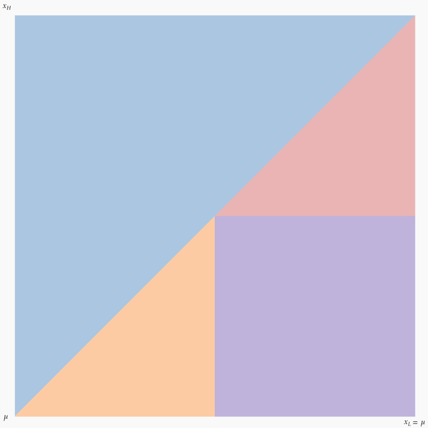
Entropy Reduction Cost: Low O.O.



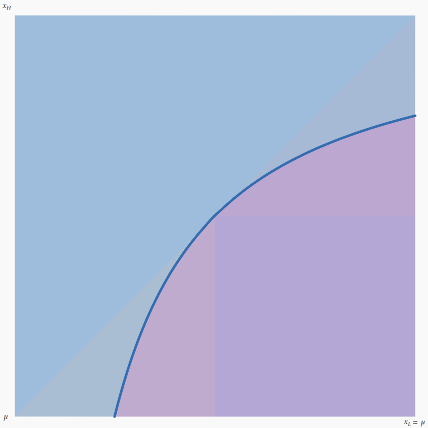
Entropy Reduction Cost: High O.O.



Entropy Reduction Cost: Moderate O.O. (interim IR)



Entropy Reduction Cost: Moderate O.O. (no interim IR)

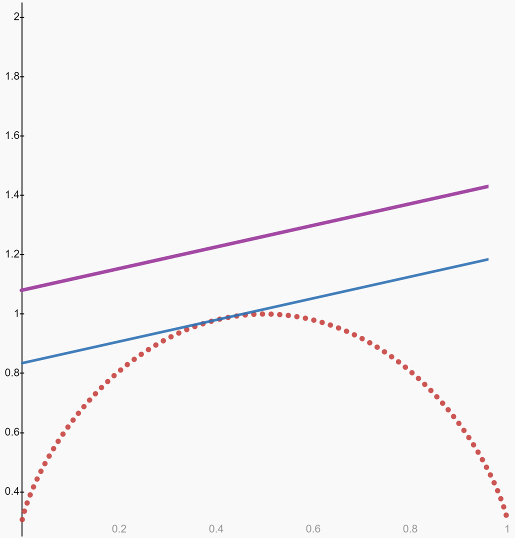


Risk-Averse Agent (& No Limited Liability)

Picking a Point on the "O.O. curve"

- With interim IR, choose a point, $\mathbf{x}^*$, on $v_0 - \kappa c(\mathbf{x})$ where $f_{\mathcal{H}}(\mathbf{x})$ is tangent.
- Generically $\mathbf{x}^* \neq \mu \Rightarrow$ Agent gets rents.
- Without interim IR, choose a "slope" of $f_{\mathcal{H}}$ that intersects (μ, v_0).
- Agent gets no rents.
- In both, efficient implementation is impossible (unless $F = \delta_\mu$).

Interim IR & RA



Discussion and Extensions

Contracts in Which the Agent Exits (w/ Positive Probability)

- Unique posterior corresponding to exit.
- For a distribution F with support on $s \leq n$ points: at most s more contracts to check (thanks to Theorem).

Recap of Findings

- P unconstrained in implementation: any F is feasible.
- An agent's optimal learning pins down the relative incentives.
- When A is risk neutral and negative transfers allowed (no l.l.), P can implement any distribution at first-best cost.
- L.l. and risk-neutral A: first-best implementation for sufficiently contracted distributions. Rents for A if first-best infeasible.
- No l.l. and risk-averse A: first-best infeasible. Rents for A (generically) with interim IR, none otherwise.

Extensions

- Prior with a density: first-best result goes through.
- More general objective: fine.

Thanks for Coming!