Lecture 2: Commitment Payoff Theorem

Long-Run Short-Run Models with Imperfect Monitoring

Harry PEI
Department of Economics, Northwestern University

Spring Quarter, 2021

Last Lecture: Reputation under Perfect Monitoring

Commitment Payoff Theorem: Fudenberg and Levine (1989)

For every $\varepsilon > 0$, there exists $T \in \mathbb{N}$,

such that when π attaches prob more than ε to commitment type $a_1^* \in \Omega^m$,

rational P1's payoff in any Bayes Nash Equilibrium is at least:

$$(1-\delta^T)\underline{u}_1 + \delta^T v_1^*(a_1^*).$$

The proof uses an elegant Bayesian learning argument:

- If rational P1 deviates and imitates commitment type a_1^*,

  then there is a uniform upper bound T on the number of periods s.t. P2 does not best reply to a_1^*.
- The upper bound T does not depend on δ.

Today: Reputation Building under Imperfect Monitoring

What happens when P2s cannot perfectly observe whether P1 has honored his commitment?

- the public signal is noisy,
- commitment action is mixed,
- extensive-form stage game, future P2s observe the terminal node,
- P1 observes an i.i.d. state before choosing his action.

Questions:

- How much payoff can a patient player guarantee?
- What is the maximal payoff a patient player can receive?

Recall: Model

- Time: $t = 0, 1, 2, \ldots$
- Long-lived player 1 (P1),
  vs an infinite sequence of short-lived player 2s (P2).

- Players move simultaneously in the stage game: $a_1 \in A_1$, $a_2 \in A_2$.
  * Actions in period t: $a_{1,t}$ and $a_{2,t}$.

- Stage-game payoffs: $u_1(a_1, a_2)$, $u_2(a_1, a_2)$.
  * P1's *discounted average payoff*: $\sum_{t=0}^{\infty}(1-\delta)\delta^t u_1(a_{1,t}, a_{2,t})$.

- Public signals: $y \in Y$, with $\rho(y|a_1, a_2)$ the probability of y.
  * y_t: public signal in period t.
  * Last lecture: $Y = A_1 \times A_2$ and $\rho(a_1, a_2|a_1, a_2) = 1$.
  * Now: general monitoring structure ρ.

Recall: Model

P1's type is perfectly persistent, draw from $\Omega \equiv \{\omega^r\} \bigcup \Omega^m$.

1. ω^r is the *rational type*.

   Can flexibly choose actions in order to maximize payoffs.

2. Each $\alpha_1^* \in \Omega^m$ represents a *commitment type*, with $\Omega^m \subset \Delta(A_1)$.

   Does not care about payoffs and plays α_1^* in every period.

P2's prior belief: $\pi \in \Delta(\Omega)$.

What can players observe?

- Player 1's history: $h_1^t \in \mathcal{H}_1^t \equiv \Omega \times \{A_1 \times A_2 \times Y\}^t$.
- Player 2's history: $h_2^t \in \mathcal{H}_2^t \equiv \{A_2 \times Y\}^t$.

Assumptions: A_1, A_2, Y and Ω^m are finite, π has full support.

What can go wrong under imperfect monitoring?

A simple example: One commitment action H.

- Suppose $Y \equiv \{G, B\}$ and $\rho(G|H) = \rho(G|L)$.

  What is player 1's equilibrium payoff when commitment prob is small?

–	T	N
H	$2, 1$	$-2, 0$
L	$3, -1$	$0, 0$

Bottomline: We need a general formula for the payoff lower bound (take the monitoring structure into account).

A More Permissive Notion of Best Reply

Let $||\cdot||$ be the total variation distance.

Definition: ε-confirming best reply

$\alpha_2 \in \Delta(A_2)$ *is an* ε*-confirming best reply to* $\alpha_1 \in \Delta(A_1)$ *if there exists* $\alpha_1' \in \Delta(A_1)$ *such that*

1. α_2 *best replies to* α_1',
2. $\left|\left|\rho(\cdot|\alpha_1, \alpha_2) - \rho(\cdot|\alpha_1', \alpha_2)\right|\right| \leq \varepsilon$.

Idea: α_2 is P2's best reply to something that is hard to distinguish from α_1.

- If y_t is uninformative, then any undominated action is an ε-confirming best reply to any $\alpha_1 \in \Delta(A_1)$.
- If y_t is more informative, then the set of ε-confirming best reply to any given α_1 shrinks.

Properties of ε-Confirming Best Reply

Let $B_\varepsilon(\alpha_1) \subset \Delta(A_2)$ be the set of P2's ε-confirming best replies to α_1.

Properties of ε-Confirming Best Reply:

1. If $\varepsilon' < \varepsilon$, then $B_{\varepsilon'}(\alpha_1) \subset B_\varepsilon(\alpha_1)$.
2. $\lim_{\varepsilon \downarrow 0} B_\varepsilon(\alpha_1) = B_0(\alpha_1)$ (why?).
3. For every $\varepsilon \geq 0$, $\mathrm{BR}_2(\alpha_1) \subset B_\varepsilon(\alpha_1)$.

Definition: Statistical Identification

P1's actions are statistically identified if for every $\alpha_2 \in \Delta(A_2)$, $\{\rho(\cdot|a_1, \alpha_2)\}_{a_1 \in A_1}$ are linearly independent vectors.

4. If P1's actions are statistically identified, then $\mathrm{BR}_2(\alpha_1) = B_0(\alpha_1), \forall \alpha_1$.

Statement of Payoff Lower Bound Result

Payoff Lower Bound Result: Fudenberg and Levine (1992)

For every $\varepsilon > 0$, there exists $\underline{\delta} \in (0,1)$

s.t. when $\delta > \underline{\delta}$ and π attaches prob more than ε to commitment type α_1^*,

rational P1's payoff in any BNE is at least:

$$\min_{\alpha_2 \in B_\varepsilon(\alpha_1^*)} u_1(\alpha_1^*, \alpha_2) - \varepsilon.$$

1. Fix the type distribution and let $\delta \to 1$, P1's payoff lower bound is:

$$\lim_{\varepsilon \downarrow 0} \min_{\alpha_2 \in B_\varepsilon(\alpha_1^*)} u_1(\alpha_1^*, \alpha_2) = \min_{a_2 \in B_0(\alpha_1^*)} u_1(\alpha_1^*, \alpha_2).$$

2. When P1's actions are statistically identified, the value of the red equation equals P1's commitment payoff from α_1^*, namely,

$$\min_{\alpha_2 \in \mathrm{BR}_2(\alpha_1^*)} u_1(\alpha_1^*, \alpha_2).$$

Proof of Payoff Lower Bound Result

Three approaches:

1. Fudenberg and Levine (1992): Doob's upcrossing inequality.
2. Sorin (1999): merging between prob measures (Blackwell and Dubins).
3. Gossner (2011): relative entropy.

Detour: Relative Entropy

Let X be a countable set, and let $p, q \in \Delta(X)$.

Relative entropy/KL-divergence of q with respect to p:

$$d(p||q) \equiv \sum_{x \in X} p(x) \log \frac{p(x)}{q(x)}.$$

Intuitively, it measures an observer's expected error in predicting $x \in X$ using the distribution q when the true distribution is p.

Thought experiment: suppose we have n i.i.d. draws from X with true distribution p but an observer's believed distribution is q.

- Log likelihood ratio of a given sample is $\sum_{x \in X} n_x \log \frac{p(x)}{q(x)}$.
- As $n \to \infty$, average log likelihood ratio goes to $d(p||q)$.

Why should we care about entropy?

Suppose σ is the equilibrium being played, but player 1 deviates and plays the equilibrium strategy of type ω.

- Let $P_\sigma \in \Delta\{Y \times A_2\}^\infty$ be the distribution over player 2s' observations in the entire game.
- Let $P_{\omega,\sigma} \in \Delta\{Y \times A_2\}^\infty$ be the distribution over player 2s' observations conditional on knowing that player 1's type is ω.

Player 2's predictions may have some errors.

- However, her prediction errors of the entire game must be bounded.

Why? $P_\sigma = \sum_{\omega \in \Omega} \pi(\omega) P_{\omega,\sigma}$, which implies that:

$$d\Big(P_{\omega,\sigma} \Big\| P_\sigma\Big) \leq \underbrace{-\log \pi(\omega)}_{\text{a bounded number}} .$$

One-Step Ahead Prediction Error

Bounding the prediction error in the entire game is not that useful.

- What matters for P2's incentives is her prediction in each period.

P2's best reply problem at history h_2^t:

- She has some belief about how P1 behaves in period t, say $\alpha_1(h_2^t)$.

  She plays a best reply to $\alpha_1(h_2^t)$.
- $\alpha_1(h_2^t)$ and this best reply induce $p_{\sigma|h_2^t} \in \Delta(Y)$.
- Let $p_{\omega,\sigma|h_2^t} \in \Delta(Y)$ be the signal distribution induced by type ω.
- If $||p_{\omega,\sigma|h_2^t} - p_{\sigma|h_2^t}|| \leq \varepsilon$, then P2 plays an ε-confirming best reply to type ω's action at h_2^t.

Two problems:

1. We only know the total prediction error, not the one in each period.
2. We need to convert relative entropy to total variation distance.

Chain Rule

Let X and Y be two sets and let $p, q \in \Delta(X \times Y)$.

Let p_X, q_X, p_Y, q_Y be the marginal distributions on X and Y.

Chain rule:

$$d(p||q) = d(p_X||q_X) + \mathbb{E}_{p_X}\left[d\left(p_Y(\cdot|x)\middle\|q_Y(\cdot|x)\right)\right].$$

How to apply this:

- Partition h_2^∞ into $\cup_{t=0}^{+\infty}\{a_{2,t}, y_t\}$.

Apply the chain rule iteratively, we have

$$-\log\pi(\omega) \geq d\left(P_{\omega,\sigma}\middle\|P_\sigma\right) = \sum_{t=0}^{\infty}\mathbb{E}_{P_{\omega,\sigma}}\Big[\underbrace{d\left(p_{\omega,\sigma|h_2^t}\middle\|p_{\sigma|h_2^t}\right)}_{\text{1-step-ahead prediction error}}\Big].$$

Pinsker's Inequality

An inequality that connects relative entropy with total variation distance:

$$||p - q|| \leq \sqrt{2d(p||q)}.$$

Implication: If $d(p||q) \leq \varepsilon^2/2$, then $||p - q|| \leq \varepsilon$.

Putting Things Together

If

$$d\left(p_{\omega,\sigma|h_2^t}\middle\|p_{\sigma|h_2^t}\right) \leq \frac{\varepsilon^2}{2},$$

then

$$\left\|p_{\omega,\sigma|h_2^t} - p_{\sigma|h_2^t}\right\| \leq \varepsilon,$$

and player 2 will play an ε-confirming best reply to type ω's action at h_2^t.

Since

$$\sum_{t=0}^{\infty} \mathbb{E}_{P_{\omega,\sigma}}\left[d\left(p_{\omega,\sigma|h_2^t}\middle\|p_{\sigma|h_2^t}\right)\right] \leq -\log \pi(\omega),$$

the expected number of periods in which $d\left(p_{\omega,\sigma|h_2^t}\middle\|p_{\sigma|h_2^t}\right) \geq \frac{\varepsilon^2}{2}$ is no more than:

$$\overline{T}(\varepsilon,\omega) \equiv \left\lceil -\frac{2\log \pi(\omega)}{\varepsilon^2}\right\rceil.$$

To Conclude the Proof

Let ω be commitment type α_1^*.

If rational P1 imitates commitment type α_1^*, then

1. In periods where $d\left(p_{\alpha_1^*,\sigma|h_2^t} \| p_{\sigma|h_2^t}\right) \leq \frac{\varepsilon^2}{2}$, P1's stage-game payoff $\geq \min_{\alpha_2 \in B_\varepsilon(\alpha_1^*)} u_1(\alpha_1^*, \alpha_2)$.
2. In expectation, there can be at most $\overline{T}(\varepsilon, \alpha_1^*)$ periods in which $d\left(p_{\alpha_1^*,\sigma|h_2^t} \| p_{\sigma|h_2^t}\right) > \frac{\varepsilon^2}{2}$.

In expectation, rational P1's payoff by imitating commitment type α_1^* is at least:

$$(1-\delta^{\overline{T}(\varepsilon,\alpha_1^*)})\underline{u}_1 + \delta^{\overline{T}(\varepsilon,\alpha_1^*)} \min_{\alpha_2 \in B_\varepsilon(\alpha_1^*)} u_1(\alpha_1^*, \alpha_2).$$

This lower bound converges to $\min_{\alpha_2 \in B_\varepsilon(\alpha_1^*)} u_1(\alpha_1^*, \alpha_2)$ as $\delta \to 1$.

Payoff Upper Bound

Apply the above argument on the rational type's equilibrium strategy:

$$d\left(P_{\omega^r,\sigma} \middle\| P_\sigma\right) = \underbrace{\sum_{t=0}^{\infty} \mathbb{E}_{P_{\omega^r,\sigma}}\left[d\left(p_{\omega^r,\sigma|h_2^t} \middle\| p_{\sigma|h_2^t}\right)\right]}_{\text{expected sum of prediction error under the strategy of type } \omega^r} \leq -\log \pi(\omega^r).$$

Payoff Upper Bound Result: Fudenberg and Levine (1992)

For every $\varepsilon > 0$, there exists $\underline{\delta} \in (0,1)$

s.t. when $\delta > \underline{\delta}$ and π attaches prob more than ε to the rational type,

rational P1's payoff in any BNE is at most:

$$\sup_{\alpha_1 \in \Delta(A_1)} \max_{\alpha_2 \in B_\varepsilon(\alpha_1)} u_1(\alpha_1, \alpha_2) + \varepsilon.$$

Payoff Lower Bound & Upper Bound

Payoff lower bound for a patient player 1:

$$\max_{\alpha_1^* \in \Omega^m} \left\{ \min_{\alpha_2 \in B_0(\alpha_1^*)} u_1(\alpha_1^*, \alpha_2) \right\}.$$

Payoff upper bound for a patient player 1:

$$\sup_{\alpha_1 \in \Delta(A_1)} \left\{ \max_{\alpha_2 \in B_0(\alpha_1)} u_1(\alpha_1, \alpha_2) \right\}.$$

If actions are identified, and Ω^m is rich enough, then under generic (u_1, u_2),

- Both the lower bound and the upper bound converge to P1's (mixed) Stackelberg payoff.

Reputation leads to a sharp prediction on patient player's equilibrium payoff.

Payoff Lower and Upper Bounds in Product Choice Game

A firm (P1) and a sequence of consumers (P2s).

–	T	N
H	$2, 1$	$-1, 0$
L	$3, -1$	$0, 0$

If there exists a commitment type that plays $(\frac{1}{2} + \varepsilon)H + (\frac{1}{2} - \varepsilon)L$, then

- patient P1's payoff lower bound is close to $\frac{5}{2} - \varepsilon$.

Patient P1's payoff upper bound is close to $\frac{5}{2}$.

Patient P1's payoff is close to $5/2$ in *all equilibria*.

This is not an equilibrium refinement. (Why?)

Caveat: Lower and Upper Bounds

Lower and upper bounds are informative only when

- P1's discount factor is close to 1.
- Actions are statistically identified.

Does not provide tight payoff bounds when δ is bounded away from 1.

- Significant weight is put on $\underline{u}_1$ in the payoff lower bound, and is put on $\overline{u}_1$ in the payoff upper bound.
- P1 can steal info rent in the short run.

Payoff bounds is uninformative when actions are not identified.

- Ely and Välimäki (03): All equilibria attain the trivial lower bound.
- Example s.t. all equilibria attain the trivial upper bound?

Tighten the Payoff Lower Bound

Our proof of the Fudenberg-Levine payoff lower bound:

1. Upper bound on the sum of divergence:

$$\sum_{t=1}^{\infty} \mathbb{E}_{p_{\alpha_1^*}}\left[d\left(p_{\alpha_1^*,\sigma|h_2^t}\middle\|p_{\sigma|h_2^t}\right)\right] \leq -\log \pi(\alpha_1^*).$$

2. When $d\left(p_{\alpha_1^*|h_2^t}\middle\|p_{\sigma|h_2^t}\right) \leq \frac{\varepsilon^2}{2}$, P2 plays an ε-confirmed best reply.

3. Expected number of periods s.t. $d\left(p_{\alpha_1^*|h_2^t}\middle\|p_{\sigma|h_2^t}\right) > \frac{\varepsilon^2}{2}$ is at most

$$\overline{T}(\varepsilon,\omega) \equiv \left\lceil -\frac{2\log \pi(\omega)}{\varepsilon^2}\right\rceil.$$

Can we further tighten this bound?

ε-Entropy Confirming Best Reply

Definition: ε-entropy confirming best reply

α_2 is an ε-entropy confirming best reply to α_1 if $\exists\, \alpha_1' \in \Delta(A_1)$ s.t.

1. $\alpha_2 \in BR_2(\alpha_1')$.
2. $d\Big(\rho(\cdot|\alpha_1, \alpha_2)\Big\|\rho(\cdot|\alpha_1', \alpha_2)\Big) \leq \varepsilon$.

Let $B_\varepsilon^e(\alpha_1)$ be the set of ε-entropy confirming best replies against α_1.

Pinkser's inequality:

$$d(P||Q) \geq 2||P-Q||^2.$$

Connections:

- ε-entropy confirming best reply $\Rightarrow \sqrt{\varepsilon/2}$-confirming best reply.

Set of entropy confirming best reply is smaller, leading to tighter bounds.

Payoff Lower Bound

Let

$$v_{\alpha_1^*}(\varepsilon) \equiv \min_{\alpha_2 \in B_\varepsilon^e(\alpha_1^*)} u_1(\alpha_1^*, \alpha_2).$$

P1's worst payoff when he plays α_1^* and P2 plays an ε-entropy confirming best reply to α_1^*. (This is a decreasing function)

If $d\left(p_{\alpha_1^*} \middle\| p_{\sigma|h_2^t}\right) \leq \varepsilon$, then P2's action $\in B_\varepsilon^e(\alpha_1^*)$ and P1's payoff $\geq v_{\alpha_1^*}(\varepsilon)$.

Let

$$\varepsilon(h_2^t) \equiv d\left(p_{\alpha_1^*} \middle\| p_{\sigma|h_2^t}\right).$$

By playing α_1^* in every period, P1's payoff is bounded from below by:

$$\mathbb{E}_{p_{\alpha_1^*}}\left[(1-\delta)\sum_{t=0}^{\infty} \delta^t v_{\alpha_1^*}\left(\varepsilon(h_2^t)\right)\right]$$

Minmax Problem

Think about the problem faced by adverse nature who chooses $\{\varepsilon(h_2^t)\}_{h_2^t \in \mathcal{H}_2}$ in order to minimize:

$$\mathbb{E}_{p_{\alpha_1^*}}\left[(1-\delta)\sum_{t=0}^{\infty}\delta^t v_{\alpha_1^*}\left(\varepsilon(h_2^t)\right)\right],$$

subject to a budget constraint on

$$\mathbb{E}_{p_{\alpha_1^*}}\left[(1-\delta)\sum_{t=0}^{\infty}\delta^t \varepsilon(h_2^t)\right] \leq ???$$

Budget Constraint

What is the upper bound on $\mathbb{E}_{p_{\alpha_1^*}}\left[(1-\delta)\sum_{t=0}^{\infty}\delta^t\varepsilon(h_2^t)\right]$?

- we know that $\sum_{t=0}^{T}\varepsilon(h_2^t) \leq -\log\pi(\alpha_1^*)$ for every $T \in \mathbb{N}\cup\{\infty\}$.

For any bounded sequence $\{x_t\}_{t\in\mathbb{N}}$, summation by parts gives

$$\sum_{t=0}^{\infty}\delta^t x_t = (1-\delta)\sum_{t=0}^{\infty}\delta^t\sum_{s=0}^{t}x_s.$$

Since $\sum_{t=0}^{T}\varepsilon(h_2^t) \leq -\log\pi(\alpha_1^*)$ for every $T \in \mathbb{N}$, we have

$$\mathbb{E}_{p_{\alpha_1^*}}\left[(1-\delta)\sum_{t=0}^{\infty}\delta^t\varepsilon(h_2^t)\right] \leq -(1-\delta)\log\pi(\alpha_1^*).$$

Minmax Problem

Think about the problem faced by adverse nature who chooses $\{\varepsilon(h_2^t)\}_{h_2^t \in \mathcal{H}_2}$ in order to minimize:

$$\mathbb{E}_{p_{\alpha_1^*}}\left[(1-\delta)\sum_{t=0}^{\infty}\delta^t v_{\alpha_1^*}\left(\varepsilon(h_2^t)\right)\right],$$

subject to a budget constraint on

$$\mathbb{E}_{p_{\alpha_1^*}}\left[(1-\delta)\sum_{t=0}^{\infty}\delta^t \varepsilon(h_2^t)\right] \leq -(1-\delta)\log \pi(\alpha_1^*).$$

Let $\overline{V}_{\alpha_1^*}(\cdot)$ be the largest convex function below $v_{\alpha_1^*}(\cdot)$, the value of the constrained minimization problem is at least:

$$\overline{V}_{\alpha_1^*}\Big(-(1-\delta)\log \pi(\alpha_1^*)\Big).$$

This gives a refined lower bound on P1's equilibrium payoff.

Payoff Upper Bound

Let

$$w_{\alpha_1^*}(\varepsilon) \equiv \max_{\alpha_2 \in B_\varepsilon^e(\alpha_1^*)} u_1(\alpha_1^*, \alpha_2).$$

In words, P1's best payoff when he plays α_1^* and P2 plays an ε-entropy confirming best reply to α_1^*.

Let $\underline{W}_{\alpha_1^*}(\cdot)$ be the smallest concave function below $w_{\alpha_1^*}(\cdot)$, player 1's payoff is bounded from above by:

$$\underline{W}_{\alpha_1^*}\Big(-(1-\delta)\log \pi(\alpha_1^*)\Big).$$

Thursday

- Pedro will present Faingold (2020): How to generalize the payoff bounds to environments with frequent interactions. (FL bound leads to uninformative answers but the refined bounds lead to sharp predictions)
- I will talk about Ely-Valimaki (2003): Due to lack-of-identification, FL's payoff lower bound is trivial. Yet there are examples in which all equilibria attain this trivial lower bound.