Lecture 1: Reputation Effects and the Commitment Payoff Theorem

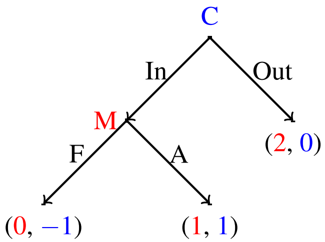
Harry PEI
Department of Economics, Northwestern University

Mini Course at Oxford University

The Chainstore Paradox

- A monopolist has branches in $T \in \mathbb{N}$ locations, with T finite.

  He faces *one potential competitor in each location*.
- In period $s \in \{1, 2, ..., T\}$, the monopolist plays against the competitor in the s-th location.



- Monopolist's total payoff is the sum of payoffs in T locations.
- Every competitor perfectly observes all actions chosen before.

The Chainstore Paradox

There is a unique subgame perfect equilibrium:

- Every competitor chooses *In* and monopolist chooses *Accommodate*.

What is wrong with this prediction?

- No matter how long the time horizon is, the monopolist never fights.
- Even if a competitor observes the monopolist fighting the past 1000 entrants, he still believes that he will be accommodated with prob 1.

Something is missing in complete information game repeated games.

Intellectual History: Commitment Type Models

How to fix this? *Gang of four*.

- Kreps and Wilson (1982), Milgrom and Roberts (1982).

Idea: Perturb the game with a small prob of commitment type.

- With probability $\varepsilon > 0$, the monopolist is irrational, doesn't care about payoffs, and mechanically fights in every period.
- With probability $1 - \varepsilon$, the monopolist is *rational*, maximizes the sum of his payoffs across periods.

Result: Gang of Four

Theorem: Gang of Four

For every $\varepsilon > 0$, there exists $T^* \in \mathbb{N}$ such that if $T \geq T^*$,

then on the equilibrium path of every sequential equilibrium,

- *The rational monopolist chooses F & each potential entrant chooses Out in all except for the last T^* periods*

Proof: Backward induction.

Takeaway: The option to build reputations can dramatically affect patient players' incentives and behaviors.

Robustness of the Gang of Four Insight?

The gang of four result relies on:

- Finite horizon and backward induction.
- Particular stage-game payoff functions.
- Entrants can perfectly observe the monopolist's action.

Another concern: Does it rely on the specification of incomplete info?

- Part 2 of Fudenberg and Maskin (1986).

Part 2 of Fudenberg and Maskin (1986)

- Let $G = (N, A, u)$ be an n-player normal form game.
- Let $\alpha^* \in \times_{i=1}^{n} \Delta(A_i)$ be a stage-game NE with payoff $\mathbf{w} \in \mathbb{R}^n$.

Folk Theorem under Incomplete Information: Fudenberg and Maskin (1986)

For any $\varepsilon > 0$ and any feasible payoff $\mathbf{v} > \mathbf{w}$, there exists $T^* \in \mathbb{N}$ such that for any $T > T^*$, there exists a strategy profile $\{s_i\}_{i \in N}$ such that in the T-fold repetition of G with public randomization where each player i is rational with probability $1 - \varepsilon$ and is committed to s_i with probability ε, there is an equilibrium where players' average payoff is within ε of $\mathbf{v}$.

Fudenberg and Levine (1989, 1992)

Extend the gang of four insights to

- environments with an infinite horizon.
- general stage game payoffs.
- imperfect monitoring.
- weaker solution concepts (Nash equilibrium).
- not sensitive to the details of incomplete info.

I will present all results in games with an infinite horizon.

- These results also apply to games with long but finite horizon.

Infinitely Repeated Game with One Long-Run Player

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

- Players simultaneously choose their actions $a_1 \in A_1$ and $a_2 \in A_2$.

  Actions in period t: $a_{1,t} \in A_1$ and $a_{2,t} \in A_2$.

- Stage-game payoffs: $u_1(a_{1,t}, a_{2,t})$, $u_2(a_{1,t}, a_{2,t})$.

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

- Public signal in period t: $y_t \in Y$,

  which is distributed according to $\rho(\cdot|a_{1,t}, a_{2,t}) \in \Delta(Y)$.

Introducing Commitment Types

P1 has persistent private info about his type $\omega \in \Omega \equiv \{\omega^r\} \bigcup \Omega^m$.

1. ω^r stands for a *rational type*, who can choose any action in order to maximize his discounted average payoff.

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

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

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

Commitment Payoff Theorem

For any commitment action $\alpha_1^* \in \Omega^m$, P1's commitment payoff from α_1^* is

$$v_1^*(\alpha_1^*) \equiv \min_{a_2 \in \mathrm{BR}_2(\alpha_1^*)} u_1(\alpha_1^*, a_2).$$

Let $\underline{u}_1$ be P1's lowest stage-game payoff.

Commitment Payoff Theorem

Suppose the monitoring technology $\rho(\cdot|a_{1,t}, a_{2,t})$ satisfies some condition.

For every $\varepsilon > 0$, there exists $\delta^* \in (0,1)$ such that when $\delta > \delta^*$ and π assigns prob more than ε to commitment type $\alpha_1^* \in \Omega^m$, the rational type of P1's payoff in any equilibrium is at least $v_1^*(\alpha_1^*) - \varepsilon$.

Commitment Payoff Theorem: Perfect Monitoring

One case in which the result applies: Perfect monitoring.

Suppose there exists a pure commitment action $a_1^* \in \Omega^m$ and the monitoring technology satisfies $Y = A_1 \times A_2$ and $\rho(a_1, a_2 | a_1, a_2) = 1$.

Commitment Payoff Theorem

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

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

the rational-type P1's payoff in any equilibrium is at least:

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

This payoff lower bound does not depend on the details of the type space.

- It only requires commitment type a_1^* to occur with positive prob.

Proof: Overview

Fix the parameters (π, δ) and an equilibrium (σ_1, σ_2).

- Consider the rational type of P1's payoff

  if he deviates from σ_1 and mechanically plays a_1^* in every period.

- Let this payoff be U_1^*.
- By definition, the rational type of P1's equilibrium payoff $\geq U_1^*$.

Proof: P1's payoff if he deviates and plays a_1^*

In every period,

- *either* P2's action is supported in $\mathrm{BR}_2(a_1^*)$.

  or P2 has an incentive to play actions outside $\mathrm{BR}_2(a_1^*)$.

In the 1st case, P1's stage-game payoff $\geq v_1^*(a_1^*)$.

In the 2nd case, there exists $\gamma > 0$ such that:

- P2 believes that a_1^* is played with prob less than $1-\gamma$ in that period.

  Such γ *depends only* on players' stage-game payoff functions.
- After P2 observes P1 plays a_1^* in that period, Bayes Rule suggests that:

$$\text{Posterior Prob of Type } a_1^* = \frac{(\text{Prior Prob of Type } a_1^*) \cdot \Pr(a_1^* | \text{type } a_1^*)}{\text{unconditional prob of } a_1^*}$$

$$\geq \frac{\text{Prior Prob of Type } a_1^*}{1-\gamma}.$$

- This can happen in at most $T \equiv \lceil \log \varepsilon / \log(1-\gamma) \rceil$ periods.

Proof: Wrap up

What is rational P1's payoff if he deviates and plays a_1^* in every period?

In periods where P2's action is supported in $\mathrm{BR}_2(a_1^*)$.

- P1's stage game payoff $\geq v_1^*(a_1^*)$.

In periods where P2's action is *not* supported in $\mathrm{BR}_2(a_1^*)$.

- P1 may receive low stage-game payoff,
- But there can be at most $T \equiv \lceil \log \varepsilon / \log(1-\gamma) \rceil$ such periods.

Lower bound on rational P1's payoff from playing a_1^* in every period:

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

This is also a lower bound for the rational-type P1's equilibrium payoff.

Some Common Misunderstandings

1. Can rational P1 convince P2s that he is a commitment type?

   Not with high prob on the equilibrium path! Belief is a martingale.

   Example: Think about a pooling equilibrium.

2. Will the rational-type P1 build a reputation?

   Not necessarily in the infinite horizon game. He may find it strictly optimal to separate from the commitment type in period 0.

3. Does it say much about the short-run players' welfare?

   No. Because rational-type P1's behavior cannot be pinned down.

Predictions on P1's Behavior?

Suppose there is a commitment type that plays P1's optimal pure commitment action a_1^* in every period, then

- What's the frequency with which the rational-type P1 plays a_1^*?

$$X^{(\sigma_1,\sigma_2)}(a_1^*) \equiv \mathbb{E}^{(\sigma_1,\sigma_2)}\left[\sum_{t=0}^{\infty}(1-\delta)\delta^t \mathbf{1}\{a_{1,t}=a_1^*\}\right]$$

Li and Pei (2021): In many games of interest, any action frequency that is compatible with

- P1 receiving payoff at least $v_1(a_1^*)$,
- P2's myopic incentives

can arise in some equilibria of the reputation game.

Li and Pei (2021)'s Theorem

Assumptions on stage-game payoffs:

- P1 has a unique optimal commitment action a_1^* and $\mathrm{BR}_2(a_1^*) = \{a_2^*\}$.
- $a_1^* \notin \mathrm{BR}_1(a_2^*)$.
- $u_1(a_1^*, a_2^*) > v^{\min} \equiv \min_{\alpha_2 \in \mathcal{A}_2} \max_{a_1 \in A_1} u_1(a_1, \alpha_2)$.

Let

$$F^*(u_1, u_2) \equiv \min_{(\alpha_1', \alpha_1'', a_2', a_2'', q) \in \Delta(A_1) \times \Delta(A_1) \times A_2 \times A_2 \times [0,1]} \left\{ q\alpha_1'(a_1^*) + (1-q)\alpha_1''(a_1^*) \right\},$$

subject to $a_2' \in \mathrm{BR}_2(\alpha_1')$, $a_2'' \in \mathrm{BR}_2(\alpha_1'')$, and

$$q u_1(\alpha_1', a_2') + (1-q) u_1(\alpha_1'', a_2'') \geq u_1(a_1^*, a_2^*).$$

Theorem: When δ is close enough to 1, rational-type P1's discounted frequency of playing a_1^* can be anything between $F^*(u_1, u_2)$ and 1.

From Perfect Monitoring to Imperfect Monitoring

Imperfect monitoring:

- The public signal is noisy.
- The commitment action is mixed.
- Extensive-form stage game and only the terminal node is observed.
- The long-run player privately observes an i.i.d. state.

Questions:

- Do we still have the commitment payoff theorem?
- How does the monitoring structure affect the patient player's payoff?

What can go wrong under imperfect monitoring?

A simple example:

- Players' stage-game payoffs:

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

- One commitment type, playing H in every period.
- Suppose $\rho(\cdot|H) = \rho(\cdot|L)$.

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

Lesson: P1's payoff depends on the monitoring technology.

A More Permissive Notion of Best Reply

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

- If $f, g \in \Delta(X)$, then $||f - g|| \equiv \frac{1}{2} \sum_{x \in X} |f(x) - g(x)|$.

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. $\left\| \rho(\cdot|\alpha_1, \alpha_2) - \rho(\cdot|\alpha_1', \alpha_2) \right\| \leq \varepsilon$.

Idea: α_2 best replies to something that is hard to distinguish from α_1.

Properties of ε-Confirming Best Reply

Let $B_\varepsilon(\alpha_1) \subset \Delta(A_2)$ denote the set of 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. $\lim_{\varepsilon \downarrow 0} B_\varepsilon(\alpha_1) = B_0(\alpha_1)$. (convince yourself)
3. $\mathrm{BR}_2(\alpha_1) \subset B_0(\alpha_1)$.

When is $B_0(\alpha_1) \subset \mathrm{BR}_2(\alpha_1)$?

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 $\mathrm{BR}_2(\alpha_1) = B_0(\alpha_1)$.
    - Why? $\rho(\cdot|\alpha_1, \alpha_2) \neq \rho(\cdot|\alpha_1', \alpha_2)$ if $\alpha_1 \neq \alpha_1'$.

Statement of the Payoff Lower Bound Result

Fudenberg and Levine (1992): Payoff Lower Bound

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

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

the rational type of P1's payoff in any equilibrium is at least:

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

1. Fix $\pi \in \Delta(\Omega)$ and let $\delta \to 1$, P1's payoff lower bound is:

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

2. When P1's actions are statistically identified,

$$\min_{\alpha_2 \in B_0(\alpha_1^*)} u_1(\alpha_1^*, \alpha_2) = \min_{\alpha_2 \in \mathrm{BR}_2(\alpha_1^*)} u_1(\alpha_1^*, \alpha_2).$$

Statement of the Payoff Upper Bound Result

Fudenberg and Levine (1992): Payoff Upper Bound

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

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

the rational type of P1's payoff in any equilibrium 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 & Payoff Upper Bound

Payoff lower bound as $\delta \to 1$:

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

Payoff upper bound as $\delta \to 1$:

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

If actions are identified and Ω^m contains the optimal commitment action,

- Both bounds converge to P1's (mixed) Stackelberg payoff.

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

Proof: Overview

Let us examine the rational type's payoff once he deviates to type ω's equilibrium strategy.

- Type ω could be a commitment type and could be a rational type.

P2's Best Reply Problem at Any Given History

Fix an equilibrium σ and consider P2's best reply problem at h_2^t:

- P2 best replies to her belief about P1's period t action $\alpha_1(h_2^t)$.
- $\alpha_1(h_2^t)$ and P2's action at h_2^t induce a distribution of y_t: $p_{h_2^t} \in \Delta(Y)$.

  This is P2's belief about y_t in period t.
- Let $p_{\omega|h_2^t} \in \Delta(Y)$ be the distribution of y_t conditional on type ω.
- If $||p_{\omega|h_2^t} - p_{h_2^t}|| \leq \varepsilon$, then P2 plays an ε-confirming best reply to type ω's equilibrium action at h_2^t.

Question: Suppose the rational type of P1 deviates and uses type ω's equilibrium strategy, then in how many periods can we have

$$||p_{\omega|h_2^t} - p_{h_2^t}|| > \varepsilon$$

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, $d(p||q)$ measures an observer's expected error in predicting $x \in X$ when he thinks that the distribution of x is q while the true distribution is p.

Proof for $d(p||q) \geq 0$: Uses the fact that $\log a \leq a - 1$ for every $a > 0$.

Why should we care about entropy?

Fix the equilibrium being played.

- Let $P \in \Delta(Y^\infty)$ be the distribution over player 2's observations.

Suppose player 1 deviates and plays the equilibrium strategy of type ω.

- Let $P_\omega \in \Delta(Y^\infty)$ be the distribution over player 2s' observations.

Since $P = \sum_{\omega \in \Omega} \pi(\omega) P_\omega$, we have

$$d\left(P_\omega \middle\| P\right) \leq \underbrace{-\log \pi(\omega)}_{\text{a bounded number}} .$$

However, we need to bound $||p_{\omega|h_2^t} - p_{h_2^t}||$, not $d\left(P_\omega \middle\| P\right)$.

1. We need to relate $d\left(P_\omega \middle\| P\right)$ to $d\left(p_{\omega|h_2^t} \middle\| p_{h_2^t}\right)$.
2. We need to convert relative entropy to total variation distance.

Chain Rule: Relate $d\left(P_\omega \middle\| P\right)$ to $d\left(p_{\omega|h_2^t} \middle\| p_{h_2^t}\right)$

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:

- h_2^∞ consists of the signal player 2 observes in each period.

Iteratively applying the chain rule, we can obtain that

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

Pinsker's Inequality: Relate Entropy to TV Distance

Pinsker's Inequality:

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

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

Putting Things Together

By Pinsker's Inequality, if

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

then

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

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

Since

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

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

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

Proof: Payoff Lower Bound

Let ω be commitment type α_1^*.

If the rational type deviates and imitates commitment type α_1^*, then

1. In periods where $d\left(p_{\alpha_1^*|h_2^t} \middle\| p_{h_2^t}\right) \leq \frac{\varepsilon^2}{2}$, 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 $d\left(p_{\alpha_1^*|h_2^t} \middle\| p_{h_2^t}\right) \geq \frac{\varepsilon^2}{2}$.

In expectation, the rational type's payoff by playing α_1^* in every period 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 $\delta \to 1$.

Proof: Payoff Upper Bound

Let ω be the rational type.

If the rational type plays his equilibrium strategy, then

1. In periods where $d\left(p_{\omega|h_2^t} \middle\| p_{h_2^t}\right) \leq \frac{\varepsilon^2}{2}$, P1's stage-game payoff is no more than $\sup_{\alpha_1 \in \Delta(A_1)} \max_{\alpha_2 \in B_\varepsilon(\alpha_1)} u_1(\alpha_1, \alpha_2)$.
2. In expectation, there can be at most $\overline{T}(\varepsilon, \omega)$ periods in which $d\left(p_{\omega|h_2^t} \middle\| p_{h_2^t}\right) \geq \frac{\varepsilon^2}{2}$.

In expectation, the rational type's payoff by playing his equilibrium strategy is at most

$$(1 - \delta^{\overline{T}(\varepsilon,\omega)})\overline{u}_1 + \delta^{\overline{T}(\varepsilon,\omega)} \sup_{\alpha_1 \in \Delta(A_1)} \max_{\alpha_2 \in B_\varepsilon(\alpha_1)} u_1(\alpha_1, \alpha_2),$$

which converges to $\sup_{\alpha_1 \in \Delta(A_1)} \max_{\alpha_2 \in B_\varepsilon(\alpha_1)} u_1(\alpha_1, \alpha_2)$ as $\delta \to 1$.