Lecture 4: Sustainability of Reputations

Harry PEI
Department of Economics, Northwestern University

Spring Quarter, 2021

Lectures 1-3: Patient Player's Payoffs

If actions are identified & positive prob of commitment type α_1^*, then

- patient player's payoff is at least his commitment payoff from α_1^*.

Proof idea:

- Patient player's payoff if he deviates and plays α_1^* in every period.
- The expected number of periods that P2 does not best reply to α_1^* is uniformly bounded from above.
- This provides a lower bound for rational P1's equilibrium payoff.

What Fudenberg and Levine's results do not tell us...

1. the patient player's equilibrium behavior,
2. the uninformed players' learning.

Perfect Monitoring: Anything Goes

Recall the characterization theorem in Li and Pei (2021).

If the informed player's actions can be perfectly monitored, then:

- $\exists$ equilibrium s.t. rational type imitates the commitment type.
- $\exists$ equilibrium s.t. imitating commitment type is strictly sub-optimal.

Imperfect Monitoring: Cripps, Mailath and Samuelson

Informal Illustration

Under some conditions on commitment action and monitoring technology, then in all Bayes Nash Equilibria,

1. *Player 2s almost surely learn player 1's type as $t \to \infty$.*

   i.e., rational P1 loses his reputation in all equilibria.
2. *Rational type will not pool with the commitment type in the long run.*

Lots of follow-up works:

- Generalize this theorem: Cripps, Mailath and Samuelson (07).
- Many ways to break this result: Hörner (02), Phelan (06), Ekmekci (11), Liu (11), Ekmekci, Gossner and Wilson (12).

Recall: Model Setup

- Time: $t = 0, 1, 2, ...$
- Long-lived player 1 (P1) with discount δ, *vs* short-lived player 2s (P2).
- Actions: $a_1 \in A_1$ and $a_2 \in A_2$.
- Stage-game payoffs: $u_1(a_1, a_2)$, $u_2(a_1, a_2)$.
- Public signals: $y \in Y$, with $\rho(y|a_1, a_2)$ the probability of y.
- P1 has two types:
    1. rational type, denoted by ω^r.
    2. commitment type $\alpha_1^* \in \Delta(A_1)$.
- P2's prior belief: commitment type occurs with prob $\pi_0 \in (0, 1)$.
- Histories: $h_1^t \in \mathcal{H}_1^t \equiv \{A_1 \times A_2 \times Y\}^t$ and $h_2^t \in \mathcal{H}_2^t \equiv \{A_2 \times Y\}^t$.
    * Main insight extends when P1 cannot observe a_2.
- **Assumptions:** A_1, A_2, Y are finite.

Conditions on Monitoring & Stage-Game Payoff

Assumption 1: Statistical Identification

For every $a_2 \in A_2$, $\{\rho(\cdot|a_1, a_2)\}_{a_1 \in A_1}$ are linearly independent.

Assumption 2: Full Support Monitoring

For every $a_2 \in A_2$ and $a_1, a_1' \in A_1$, the support of $\rho(\cdot|a_1, a_2)$ coincides with the support of $\rho(\cdot|a_1', a_2)$.

Assumption 3: Strict Best Reply & Lack of Commitment

There exists $a_2^* \in A_2$ such that $BR_2(\alpha_1^*) = \{a_2^*\}$.

(α_1^*, a_2^*) is not a stage-game Nash Equilibrium.

Assumption on unique best reply:

- satisfied under generic $\alpha_1^* \in \Delta(A_1)$ and u_2.

Disappearing Reputation Theorem

Let π_t be the prob of commitment type α_1^* in period t.

For every $\sigma \equiv (\sigma_{\omega^r}, \sigma_2)$, let $P_{\omega^r,\sigma}$ be the prob measure over histories when:

- P1 plays the rational type's equilibrium strategy σ_{ω^r} and P2s play σ_2.

Disapprearing Reputation Theorem

Under Assumptions 1, 2, and 3, in every BNE σ:

$$\lim_{t \to \infty} \pi_t = 0, \quad P_{\omega^r,\sigma} \text{ almost surely.}$$

Bottomline: P2s learn P1's type with probability 1.

- Applies to every $(\delta, \pi_0) \in (0,1)^2$ and every equilibrium.
- It is an asymptotic result, i.e., applies only to $t \to \infty$.

  No claim on how the rate of convergence depends on the parameters.

Disappearing Reputation Theorem meets FL

Does this result contradict FL's commitment payoff theorem? No.

- CMS's result is about what happens as $t \to \infty$.
- FL's result is about P1's discounted average payoff.

  If P1 plays his *equilibrium strategy*, then it could be the case that his *discounted average payoff is high*, but his payoff is low as $t \to \infty$.

Proof: An Intuitive Explanation

P1's actions are identified $\Rightarrow$ If P1 is rational but π_t does not converge to 0,

- then P2's belief about rational P1's action converges to α_1^*.

If P2s are almost sure that P1's action is close to α_1^* in the next τ periods,

- Full support monitoring: She has a strict incentive to play a_2^* regardless of the public signals in the next τ periods.

Since P1 can identify P2's action, P1 can learn that P2's actions in the next τ periods will be irresponsive to the public signals.

At least one action in $\text{supp}(\alpha_1^*)$ is not a myopic best reply against a_2^*.

- For every δ, there exists large enough $\tau \in \mathbb{N}$ such that

  rational P1 has a strict incentive not to play α_1^*.

If P2 realizes that rational P1 has a strict incentive to deviate from α_1^*, then it contradicts P2's belief about rational P1's action being close to α_1^*.

A Result on Asymptotic Learning

Lemma

If P1's actions are identified, then in every equilibrium σ,

$$\lim_{t\to\infty} \pi_t(1-\pi_t)\left\|\alpha_1^* - \underbrace{\mathbb{E}_\sigma[\sigma_{\omega^r}(h_1^t)|h_2^t]}_{\text{P2's expectation of rational P1's action at } h_2^t}\right\| = 0, \quad P_\sigma \text{ a.s.}$$

Since y_t can statistically identify P1's action,

- If rational P1 behaves differently from the commitment type for unbounded number of periods,

  then P2 will learn P1's type almost surely as $t \to +\infty$.

Since player 2's belief cannot be wrong,

$$\lim_{t\to\infty} \pi_t\left\|\alpha_1^* - \mathbb{E}_\sigma[\sigma_{\omega^r}(h_1^t)|h_2^t]\right\| = 0, \quad P_{\omega^r,\sigma} \text{ a.s.}$$

An Implication of Sustainable Reputation

Recall that

$$\lim_{t\to\infty} \pi_t \left\| \alpha_1^* - \mathbb{E}_\sigma[\sigma_{\omega^r}(h_1^t)|h_2^t] \right\| = 0, \quad P_{\omega^r,\sigma} \text{ a.s.}$$

Suppose there exists a positive prob event under $P_{\omega^r,\sigma}$ such that $\pi_t \nrightarrow 0$, then for every $\varepsilon > 0$, there exists $T \in \mathbb{N}$ such that event

$$B \equiv \bigcap_{t\geq T} \left\{ \left\| \alpha_1^* - \mathbb{E}_\sigma[\sigma_{\omega^r}(h_1^t)|h_2^t] \right\| < \varepsilon \right\}$$

occurs with strictly positive probability under $P_{\omega^r,\sigma}$.

Pick ε s.t. $\{a_2^*\} = \text{BR}_2(\alpha_1)$ for every α_1 with $||\alpha_1 - \alpha_1^*|| < \varepsilon$.

- If event B happens, then P2s have strict incentives to play a_2^* starting from period T.

Recap

What we want to show to get a contradiction:

- P1 has a strict incentive not to play α_1^* when event B happens.
- P2s realize P1's strict incentive of not playing α_1^*.

  This contradicts the presumption that $\left\|\alpha_1^* - \mathbb{E}_\sigma[\sigma_{\omega^r}(h_1^t)|h_2^t]\right\| < \varepsilon$.

Two obstacles to derive such a contradiction:

1. P1 doesn't know event B in finite time.
2. P2 may not learn the event in which P1 has a strict incentive to deviate.

P1's belief about P2's future actions under *rational P1's equilibrium strategy*

$$B \equiv \bigcap_{t \geq T} \left\{ \left\| \alpha_1^* - \mathbb{E}_\sigma[\sigma_{\omega^r}(h_1^t) | h_2^t] \right\| < \varepsilon \right\}$$

Lemma

For every $\tau \in \mathbb{N}$, there exists a subsequence $\{t_n\}_{n \in \mathbb{N}}$ such that as $n \to \infty$,

$$\mathbf{1}_B \sum_{k=1}^{\tau} \left\{ 1 - \underbrace{\mathbb{E}_{P_{\omega^r,\sigma}}[\sigma_2(h_2^{t_n+k})(a_2^*) | h_1^{t_n}]}_{\text{P1's prediction about P2's future action in equilibrium}} \right\} \to 0, \quad P_{\omega^r,\sigma} \text{ a.s.}$$

If rational P1 plays his equilibrium strategy and event B happens,

- then P1 predicts that P2 will play a_2^* with prob close to 1 in each of the next τ periods.

Why? Let's use the argument in Lecture 2

Two probability measures:

- $P_{\omega^r,\sigma}$: P1 plays rational type's equilibrium strategy and P2s play σ_2.
- $P_{\omega^r,\sigma|B}$: ... conditional on event B.

Suppose the true DGP is $P_{\omega^r,\sigma|B}$, P1's believed DGP is $P_{\omega^r,\sigma}$.

Think about P1's prediction of $\{(a_{1,k}, a_{2,k}, y_k)\}_{k=0}^{\infty}$ in period 0:

$$-\log P_{\omega^r,\sigma}(B) \geq d\left(P_{\omega^r,\sigma|B} \middle\| P_{\omega^r,\sigma}\right) = \sum_{t=0}^{\infty} \mathbb{E}_{P_{\omega_s,\sigma|B}}\Big[\underbrace{d\left(p_{\omega^r,\sigma|B,h_1^t} \middle\| p_{\omega^r,\sigma|h_1^t}\right)}_{\text{P1's 1-step-ahead prediction error}} \Big].$$

where

- $p_{\omega^r,\sigma|h_1^t}$ is P1's prediction of $\{a_{1,t}, a_{2,t}.y_t\}$ at h_1^t.
- $p_{\omega^r,\sigma|B,h_1^t}$ is P1's prediction of $\{a_{1,t}, a_{2,t}.y_t\}$ at h_1^t conditional on B.

Why? Let's use the argument in Lecture 2

$$-\log P_{\omega^r,\sigma}(B) \geq \sum_{t=0}^{\infty} \mathbb{E}_{P_{\omega^r,\sigma|B}} \Big[\underbrace{d\Big(p_{\omega^r,\sigma|B,h_1^t} \Big\| p_{\omega^r,\sigma|h_1^t}\Big)}_{\text{P1's 1-step-ahead prediction error}} \Big].$$

implies that for any $\tau \in \mathbb{N}$,

$$\lim_{k\to\infty} \sum_{t=k}^{k+\tau} \mathbb{E}_{P_{\omega^r,\sigma|B}} \Big[\underbrace{d\Big(p_{\omega^r,\sigma|B,h_1^t} \Big\| p_{\omega^r,\sigma|h_1^t}\Big)}_{\text{P1's 1-step-ahead prediction error}} \Big] = 0.$$

$\sum_{t=k}^{k+\tau} \mathbb{E}_{P_{\omega^r,\sigma|B}} \Big[d\Big(p_{\omega^r,\sigma|B,h_1^t} \Big\| p_{\omega^r,\sigma|h_1^t}\Big) \Big] \geq$ the divergence between:

- P1's prediction of $\{a_{1,k+m}, a_{2,k+m}, y_{k+m}\}$ at h_1^k,
- P1's prediction of $\{a_{1,k+m}, a_{2,k+m}, y_{k+m}\}$ at h_1^k conditional on B.

for any $m \in \{0, 1, 2, ..., \tau\}$.

Why? Let's use the argument in Lecture 2

P1's prediction error of $\{a_{1,k+m}, a_{2,k+m}, y_{k+m}\}$ being small implies that:

- P1 expects P2 to play actions close to a_2^* in period $k+m$ at h_1^k,

$$\mathbf{1}_B \sum_{t=1}^{\tau} \Big\{ 1 - \mathbb{E}_{P_{\omega^r,\sigma}}[\sigma_2(h_2^{k+t})(a_2^*)|h_1^k] \Big\} \to 0 \text{ in prob.}$$

Convergence in prob implies convergence a.s. in a subsequence.

Convergence when P1 deviates

Previous lemma: Along a subsequence of periods,

$$\mathbf{1}_B \sum_{k=1}^{\tau} \left\{ 1 - \underbrace{\mathbb{E}_{P_{\omega^r,\sigma}}[\sigma_2(h_2^{t_n+k})(a_2^*)|h_1^t]}_{\text{P1's prediction about P2's action \textit{in equilibrium}}} \right\} \to 0 \quad P_{\omega^r,\sigma} \text{ a.s.}$$

Does it imply rational P1's incentive not to play α_1^*? No!

- It's P1's prediction about P2's action under P1's equilibrium strategy.
- Does not say what happens when P1 deviates.

For every $t, \tau \in \mathbb{N}$, let

$$B_t(\tau) \equiv \left\{ h_2^t \middle| \sigma_2(h_2^{t+k})(a_2^*) > 1 - \varepsilon, \forall h_2^{t+k},\ \forall k = 1, 2, ..., \tau \right\}$$

Full support monitoring implies that

Lemma

For every $\tau \in \mathbb{N}$, $\exists$ subsequence $\{t_n\}_{n \in \mathbb{N}}$ s.t. as $n \to \infty$,

$$P_{\omega^s,\sigma}(B_{t_n}(\tau)|h_1^{t_n})\mathbf{1}_B \to \mathbf{1}_B, \quad P_{\omega^r,\sigma}\ a.s.$$

P2s figure out P1's strict incentive not to play α_1^*

Lemma

Along this subsequence $\{t_n\}_{n\in\mathbb{N}}$,

$$\mathbb{E}_{P_{\omega^r,\sigma}}\left[P_{\omega^r,\sigma}(B_{t_n}(\tau)|h_1^{t_n})\mathbf{1}_B\middle|h_2^{t_n}\right] \to \mathbf{1}_B, \quad P_{\omega^r,\sigma}\ a.s.$$

P2 eventually figures out rational P1's incentive to deviate from α_1^*

- Why? Because P2 learns B in the long run.

Proof: From the previous lemma,

$$P_{\omega^r,\sigma}(B_{t_n}(\tau)|h_1^{t_n})\mathbf{1}_B \to \mathbf{1}_B, \quad P_{\omega^r,\sigma}\ a.s.$$

Therefore,

$$\mathbb{E}_{P_{\omega^r,\sigma}}\left[P_{\omega^r,\sigma}(B_{t_n}(\tau)|h_1^{t_n})\mathbf{1}_B\middle|h_2^{t_n}\right] \to \mathbb{E}_{P_{\omega^r,\sigma}}\left[\mathbf{1}_B\middle|h_2^{t_n}\right]$$

Since B is measurable with respect to $\mathcal{H}_2^\infty$,

$$\lim_{n\to\infty}\mathbb{E}_{P_{\omega^r,\sigma}}\left[\mathbf{1}_B\middle|h_2^{t_n}\right] = \mathbf{1}_B.$$

Wrapping Up

If $\pi_t \nrightarrow 0$, $P_{\omega^r,\sigma}$ a.s., then there exists a positive prob event B:

- P2's belief about rational P1's action converges to α_1^*.

We have also found a subsequence of periods s.t. conditional on event B,

1. P1 has a strict incentive not to play some actions in the support of α_1^*.
2. P2s asymptotically know that this will happen.

This contradicts the previous conclusion that P2's belief about rational P1's action converges to α_1^* in event B.

When are reputations sustainable?

When are reputations sustainable?

- P1's type changes over time (Cole, Dow and English 95, Phelan 06, Ekmekci, Gossner and Wilson 12).
- P2 has limited memory (Liu 11, Liu and Skrzypacz 14).
- Censor information about P1's past actions (Ekmekci 11).
- Multiple long-run players competing for customers (Hörner 02).

Ekmekci, Gossner and Wilson (2012)

Modified reputation model:

- In each period, P1 dies with prob $\rho \in (0, 1)$

  and is replaced by a new P1 with type drawn according to $\pi \in \Delta(\Omega)$.

- P2 cannot observe these replacements.

  e.g., restaurant changing chefs or ownership.

Main Results:

1. P1's payoff lower bound when types are changing over time.
2. P1's payoff lower bound at every on-path history.

Takeaway: Reputations are sustainable when types are changing.

Sustainable Reputation Theorem

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. $d\Big(\rho(\cdot|\alpha_1, \alpha_2)\Big\|\rho(\cdot|\alpha_1', \alpha_2)\Big) < \varepsilon$.

Let $v_{\alpha_1}(\varepsilon) \equiv \min_{\alpha_2 \in \mathrm{B}_\varepsilon^e(\alpha_1)} u_1(\alpha_1, \alpha_2)$ and $w_{\alpha_1}(\cdot)$ be the largest convex function below $v_{\alpha_1}(\cdot)$.

Theorem: Payoff Lower Bound with Replacement

If $\alpha_1^* \in \Omega^m$, then rational player 1's ex ante payoff in any BNE is at least:

$$w_{\alpha_1^*}\Big(-(1-\delta)\log \pi(\alpha_1^*) - \log(1-\rho)\Big)$$

Rational player 1's continuation value at any on-path history in any BNE is at least:

$$w_{\alpha_1^*}\Big(-(1-\delta)\log \rho\pi(\alpha_1^*) - \log(1-\rho)\Big).$$

Interpretation

Payoff lower bound on P1's ex ante payoff with constant type:

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

Lower bound on P1's ex ante payoff with changing types:

$$w_{\alpha_1^*}\Big(-(1-\delta)\log \pi(\alpha_1^*) - \underbrace{\log(1-\rho)}_{\text{the effects of secret replacements}}\Big)$$

Lower bound on P1's continuation payoff with changing types:

$$w_{\alpha_1^*}\Big(-(1-\delta)\log \rho\pi(\alpha_1^*) - \log(1-\rho)\Big).$$

Replacement has two effects:

1. It makes reputation building less profitable: $-\log(1-\rho)$.
2. It makes high continuation values sustainable: $\log \rho\pi(\alpha_1^*)$.

Proof

$$\sum_{\tau=0}^{t} \mathbb{E}_{P_{\alpha_1^*,\sigma}}\left[d(\rho(\cdot|\alpha_1^*)\|\rho(\cdot|h_2^\tau))\right] = \underbrace{d\left(P_{\alpha_1^*,\sigma}^t \middle\| P_\sigma^t\right)}_{\text{P2's prediction error on his observation } t \text{ periods from now}}.$$

Since the prob that P1's type is α_1^* from period 0 to period $t \geq \pi(\alpha_1^*)(1-\rho)^t$,

$$d\left(P_{\alpha_1^*,\sigma}^t \middle\| P_\sigma^t\right) \leq -\log \pi(\alpha_1^*) - t\log(1-\rho).$$

P1's ex ante payoff is at least

$$w_{\alpha_1^*}\left((1-\delta)\sum_{t=0}^{\infty} \delta^t \mathbb{E}_{P_{\alpha_1^*,\sigma}}\left[d(\rho(\cdot|\alpha_1^*)\|\rho(\cdot|h_2^t))\right]\right).$$

$$\text{red term} = (1-\delta)^2 \sum_{t=0}^{\infty} \delta^t \sum_{\tau=0}^{t} \mathbb{E}_{P_{\alpha_1^*,\sigma}}\left[d(\rho(\cdot|\alpha_1^*)\|\rho(\cdot|h_2^\tau))\right]$$

$$= (1-\delta)^2 \sum_{t=0}^{\infty} \delta^t d\left(P_{\alpha_1^*,\sigma}^t \middle\| P_\sigma^t\right) \leq -(1-\delta)\log \pi(\alpha_1^*) - \log(1-\rho).$$

Next Lecture

Reputation effects with interdependent values (Pei 2020 ECMA, 2021 working paper).

- relax the private value assumption.