12. Reputations with Short-Lived Players

ECON 7219 – Games With Incomplete Information

Benjamin Bernard

Imperfect Monitoring

Newly Opened Restaurant

Imperfect monitoring:

- Social beliefs are updated through Google/Yelp reviews.
- Those are based on the individual experiences of the customers, but they may not always accurately reflect the restaurant owner's effort:
  - The customer may not like that kind of food,
  - The customer's mood will have an impact on the review, etc.
- Because the reviews are correlated to the restaurant owner's efforts, we can still use them to update our beliefs.

University Reputation

UNIVERSITÉ DU LUXEMBOURG

University of Luxembourg:

- Founded in 2003 and wishes to build a good reputation.
- Pays professors very well, trying to attract international talents.
- The observable output is the graduating students' success.
- Quality of students and success after graduation are stochastic, but the distribution is correlated with the university's efforts.
- Will the university achieve its intended reputation?

Points We Have to Address

Payoffs:

- If players receive payoffs every period, those payoffs cannot reveal the chosen actions, otherwise it is a game of perfect monitoring.
- Each player's payoffs can depend only on that player's action and the information he/she has available. How much of a complication is this?

Beliefs and reputations:

- Is Bayesian updating significantly harder if actions are not observed?
- Is it harder/easier to build a reputation?
- Are reputations maintained forever as in the perfect monitoring case?
- Does a result like the reputation bound hold?

Public Signal in the Stage Game

Public signal:

- Let $\mathcal{Y}$ be the space of all possible signals with typical element $y \in \mathcal{Y}$.
- The public signal Y is a $\mathcal{Y}$-valued random variable, whose conditional distribution $\pi(\cdot \mid a)$ depends on the *realized* action profile $A = a$.
- Conditional on $A = a$, the distribution of Y does not depend on θ.

Distribution over outcomes:

- The outcome of the game is the random variable (θ, A, Y).
- For a strategy profile σ and common prior beliefs $\mu \in \Delta(\Theta)$, the probability measure P_σ is extended to $\Theta \times \mathcal{A} \times \mathcal{Y}$ by setting

$$P_\sigma(\theta = \vartheta, A = a, Y = y) = \pi(y \mid a)\sigma(\vartheta; a)\mu(\vartheta).$$

Payoffs in the Stage Game

Stage-game payoffs:

- Ex-post payoff $u_i(\vartheta_i, a_i, y)$ of player i depends only on a_i and y.
- Players maximize their ex-ante payoffs

$$u_i(\vartheta_i, a) := \mathbb{E}_a[u_i(\vartheta_i, A_i, Y)].$$

- Ex-ante payoffs of a mixed action profile α is well defined:

$$\begin{aligned} u_i(\vartheta_i, \alpha) := \mathbb{E}_{\vartheta_i,\alpha}[u_i(\vartheta_i, A_i, Y)] &= \sum_{a \in \mathcal{A}} \mathbb{E}_{\vartheta_i,\alpha}[u_i(\vartheta_i, a_i, Y) \mid A = a]\alpha(a) \\ &= \sum_{a \in \mathcal{A}} \sum_{y \in \mathcal{Y}} u_i(\vartheta_i, a_i, y)\pi(y \mid a)\alpha(a) \\ &= \sum_{a \in \mathcal{A}} u_i(\vartheta_i, a)\alpha(a) = \mathbb{E}_{\vartheta_i,\alpha}[u_i(\vartheta_i, A)]. \end{aligned}$$

Ex-Post Payoffs in the Product-Choice Game

	H	L
$\bar{y}$	$\frac{3(1-q)}{p-q}$	$\frac{2p-3q}{p-q}$
$\underline{y}$	$-\frac{3q}{p-q}$	$\frac{p-2q}{p-q}$

Ex-post payoffs

	H	L
H	$2, 3$	$0, 2$
L	$3, 0$	$1, 1$

Ex-ante payoffs

Customer satisfaction:

- Suppose that the customer satisfaction Y can either be high ($\bar{y}$) or low ($\underline{y}$) with conditional distribution given by

$$\pi(\bar{y} \mid a) = \begin{cases} p & \text{if } a_1 = H, \\ q & \text{if } a_1 = L, \end{cases} \qquad \pi(\underline{y} \mid a) = \begin{cases} 1-p & \text{if } a_1 = H, \\ 1-q & \text{if } a_1 = L. \end{cases}$$

  with $1 > p > q > 0$ so that customer satisfaction increases with effort.
- For the above ex-post payoffs, ex-ante payoffs are equal to last week's.

Public Signal in the Repeated Game

Definition 12.1

The public signal $Y = (Y^t)_{t\geq 0}$ is a $\mathcal{Y}$-valued stochastic sequence satisfying:

1. Y^t is independent of Y^s for any $s < t$,
2. Conditional on A^t, Y^t is independent of θ and A^s for any $s < t$,
3. Conditional on $A^t = a$, the distribution of Y^t is $\pi(\,\cdot \mid a)$.

Private histories:

- Player i's private history at time t is $h_i^t = (a_i^0, y^0, \dots, a_i^{t-1}, y^{t-1})$.
- We denote by $\mathcal{H}_i^t$ the set of i's private histories of length t and by $\mathcal{H}_i = \bigcup_{t=0}^{\infty} \mathcal{H}_i^t$ the set of all of i's private histories.
- As usual, we denote by $H_i^t = (A_i^0, Y^0, \dots, A_i^{t-1}, Y^{t-1})$ player i's random history of length t and by $H_i = (H_i^t)_{t\geq 0}$ the entire sequence.

Payoffs in the Reputation Game

Player 2's payoff:

- Given history $h^{t+1} = (h^t, a_2^t, y^t)$, player 2 updates their beliefs to

$$\mu(h^{t+1};\vartheta) = \frac{\sum_{a_1^t \in \mathcal{A}_1} \pi(y^t \mid a^t)\sigma_1(\vartheta, h^t; a_1^t)\mu(h^t;\vartheta)}{\sum_{\vartheta' \in \Theta}\sum_{a_1^t \in \mathcal{A}_1} \pi(y^t \mid a^t)\sigma(\vartheta', h^t; a^t)\mu(h^t;\vartheta')}.$$

- Player 2 maximizes ex-ante stage-game payoff given beliefs $\mu(h^t)$.

Player 1's payoff:

- Writing $u_1(\vartheta_p, a, y) = u_1(a, y)$, the payoff type ϑ_p maximizes

$$U_1(\sigma) := (1-\delta)\sum_{t=0}^{\infty}\delta^t \mathbb{E}_\sigma\left[u_1(A^t, Y^t)\right] = (1-\delta)\sum_{t=0}^{\infty}\delta^t \mathbb{E}_\sigma\left[u_1(A^t)\right],$$

where $u_1(a) = u_1(\vartheta_p, a)$ is defined on slide 5.

First Implications of Imperfect Monitoring

Imperfect monitoring in theory:

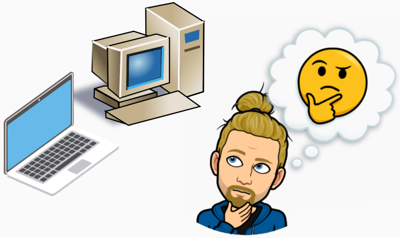
- Updating of beliefs poses no conceptual problem.
- Players' payoffs in the reputation game depend only on ex-ante stage game payoffs, hence we need not specify ex-post payoffs.

Imperfect monitoring in practice:

- Even commitment types cannot prevent a bad outcome.
- This may be an advantage or a disadvantage to player 1:
  - It may be possible gain extra utility by deviating without being noticed.
  - It may take be more difficult to build a reputation.
- How much harder is it to find Nash equilibria in a given game?

Product Choice with Imperfect Monitoring

	H	L
H	2, 3	0, 2
L	3, 0	1, 1



Product choice:

- Player 1 is either the commitment type ϑ_H or the payoff type ϑ_p.
- For $1 > p > q > 0$, the customer satisfaction Y is either high ($\bar{y}$) or low ($\underline{y}$) with conditional distribution given by

$$\pi(\bar{y} \mid a) = \begin{cases} p & \text{if } a_1 = H, \\ q & \text{if } a_1 = L, \end{cases} \qquad \pi(\underline{y} \mid a) = \begin{cases} 1-p & \text{if } a_1 = H, \\ 1-q & \text{if } a_1 = L. \end{cases}$$

- What is the Nash equilibrium in the twice-repeated game?

Parametrizing the Strategy

Parametrizing the Firm's strategy:

- In any Nash equilibrium σ, the commitment type must choose H after any history and the payoff type must choose L in the second period.
- The Firm's strategy σ_1 is entirely parametrized by $\sigma_1(\vartheta_p, \emptyset; H) = x$.

Parametrizing the Customers' strategy:

- Let z_0 denote the probability of choosing H in the first period.
- Let $\bar{z}$ and $\underline{z}$ denote the probability of choosing H after $\bar{y}$, and $\underline{y}$.

Commitment power:

- If the Firm had commitment power in the first period, we could replace z_0, $\bar{z}$, and $\underline{z}$ by $1_{\{\mu(h)\geq\frac{1}{2}\}}$ and optimize for x like we did last time.
- Without commitment power, we verify mutual best responses as usual.

Updating Beliefs

Posterior beliefs:

- Let us identify $\mu(h) \cong \mu(h; \vartheta_H)$ and write $\mu(h; \vartheta_p) = 1 - \mu$.
- $\mu_1 := \mu(Y^1)$ is a random variable, taking values

$$\mu(\bar{y}) = \frac{p\mu_0}{p\mu_0 + (xp + (1-x)q)(1-\mu_0)},$$

$$\mu(\underline{y}) = \frac{(1-p)\mu_0}{(1-p)\mu_0 + (x(1-p) + (1-x)(1-q))(1-\mu_0)}.$$

  on the events $\{Y^1 = \bar{y}\}$ and $\{Y^1 = \underline{y}\}$, respectively.
- Each of the possible posterior beliefs arises with probability

$$P_\sigma(\mu_1 = \mu(\bar{y})) = px + q(1-x),$$

$$P_\sigma\left(\mu_1 = \mu(\underline{y})\right) = (1-p)x + (1-q)(1-x).$$

Customers' Best Response

Customers' best response:

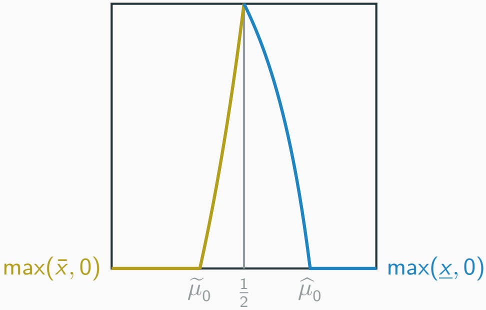
- Customers are willing to play H after history h if and only if $\mu(h) \geq \frac{1}{2}$.
- In period 0, this is equivalent to $x \geq x_0 := \frac{1-2\mu_0}{2(1-\mu_0)}$.
- After signal $\bar{y}$ and $\underline{y}$, it is equivalent to $x \leq \bar{x}$ and $x \geq \underline{x}$, where

$$\bar{x} := \frac{p\mu_0 - q(1-\mu_0)}{(p-q)(1-\mu_0)}, \qquad \underline{x} := \frac{(1-q)(1-\mu_0) - (1-p)\mu_0}{(p-q)(1-\mu_0)}.$$

- The best-response correspondence is

$$\mathcal{B}_2(x, \bar{y}) = \begin{cases} \bar{z} = 1 & \text{if } x < \bar{x}, \\ \bar{z} \in [0,1] & \text{if } x = \bar{x}, \\ \bar{z} = 0 & \text{if } x > \bar{x}, \end{cases} \qquad \mathcal{B}_2(x, \underline{y}) = \begin{cases} \underline{z} = 1 & \text{if } x > \underline{x}, \\ \underline{z} \in [0,1] & \text{if } x = \underline{x}, \\ \underline{z} = 0 & \text{if } x < \underline{x}. \end{cases}$$

Indifference Probabilities of Effort



Comparison of $\bar{x}$ and $\underline{x}$:

- It is straightforward to verify that $\bar{x}$ is increasing and $\underline{x}$ is decreasing in μ_0 and that they are equal at $\mu_0 = \frac{1}{2}$, where $\bar{x} = \underline{x} = 1$.
- The plot is for $p = 0.8$ and $q = 0.4$.

Extreme Prior Beliefs

Extreme optimism:

- If $\underline{x} \leq 0$, then the Customers play H in period 1 after any signal.
- Solving $\underline{x} \leq 0$ for μ_0 we observe that this is the case if $\mu_0 \geq \widehat{\mu}_0$, where

$$\widehat{\mu}_0 = \frac{1-q}{2-p-q} > \frac{1}{2}.$$

Extreme pessimism:

- If $\bar{x} < 0$, then the Customers play L in period 1 after any signal.
- Solving $\bar{x} < 0$ for μ_0 we observe that this is the case if $\mu_0 < \widetilde{\mu}_0$, where

$$\widetilde{\mu}_0 = \frac{q}{p+q} < \frac{1}{2}.$$

In either case: play in the first period has no effect on payoffs in the second period, hence the Firm must choose L.

Firms' Best Response

Firms' best response:

- The strategic type maximizes the discounted sum of payoffs

$$\begin{aligned} U_1(x) &= \mathbb{E}_\sigma\left[u_1(A_1^0, A_2^0)\right] + \delta \mathbb{E}_\sigma\left[u_1(L, A_2^1)\right] \\ &= (1-x) + 2z_0 + \delta + 2\delta\big(px + q(1-x)\big)\bar{z} \\ &\quad + 2\delta\big((1-p)x + (1-q)(1-x)\big)\underline{z}. \end{aligned}$$

- The partial derivative with respect to x is

$$\frac{\partial U_1(x)}{\partial x} = -1 + 2\delta(p-q)(\bar{z} - \underline{z}).$$

- The best-response correspondence is

$$\mathcal{B}_1(z_0, \bar{z}, \underline{z}) = \begin{cases} x = 1 & \text{if } 2\delta(p-q)(\bar{z}-\underline{z}) > 1, \\ x \in [0,1] & \text{if } 2\delta(p-q)(\bar{z}-\underline{z}) = 1, \\ x = 0 & \text{if } 2\delta(p-q)(\bar{z}-\underline{z}) < 1. \end{cases}$$

Nash Equilibria

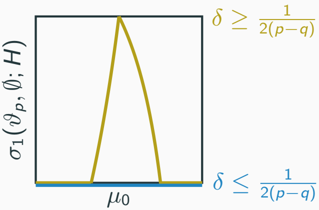
Patience:

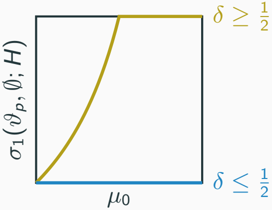
- Since $\bar{z} - \underline{z} \leq 1$, then the Firm will always play L if $\delta < \frac{1}{2(p-q)} = \delta_*$.
- Observe that δ_* is increasing in the informativeness $p - q$ of signal $\bar{y}$.
- Suppose, therefore, that $\delta > \delta_*$.

Consistency:

- Consider $\mu_0 \in \left(\widetilde{\mu}_0, \frac{1}{2}\right)$ so that $\underline{z} = 0$ for any x. Then:
  - $x = 1$ implies $\bar{z} = 0$, hence $x = 0$.
  - $x = 0$ implies $\bar{z} = 1$, hence $x = 1$.
  - $x \in [0, 1]$ implies $\bar{z} = \frac{1}{2\delta(p-q)} < 1$, hence $x = \bar{x}$.
- For $\mu_0 \in \left(\frac{1}{2}, \widehat{\mu}_0\right)$, we obtain $\bar{z} = 0$, $\underline{z} = 1 - \frac{1}{2\delta(p-q)}$, and $x = \underline{x}$.
- For $\mu = \frac{1}{2}$, we obtain $x = 1$ and $\bar{z} \geq \underline{z} + \frac{1}{2\delta(p-q)}$.

Comparison to Perfect Monitoring





Repeated play of L for patient players:

- With perfect monitoring, repeated play of L is *never* an equilibrium.
- With imperfect monitoring, it may be an equilibrium for two reasons.
- For pessimistic Customers, the reputation is not built quickly enough.
- For optimistic Customers:
  - Because deviations are not unambiguously detected, the lack of effort does not have a sufficiently strong impact on extreme beliefs.
  - The payoff type cannot resist the temptation to deviate.

Check Your Understanding

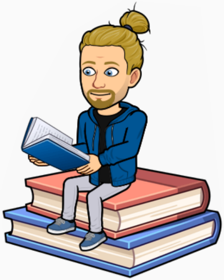
True or false:

1. With imperfect monitoring, beliefs of being a certain commitment type $\vartheta_{\widehat{a}_1}$ may decrease even if the payoff type invariably plays $\widehat{a}_1$.

Scan me

Literature

- D. Fudenberg and J. Tirole: Game Theory, Chapters 5.5–5.6, MIT Press, 1991
- G.J. Mailath and L. Samuelson: Repeated Games and Reputations: Long-Run Relationships, Chapters 7–9, Oxford University Press, 2006
- D. Abreu, D. Pearce and E. Stacchetti: Toward a theory of discounted repeated games with imperfect monitoring, Econometrica, **58** (1990), 1041–1063
- D. Fudenberg and D.K. Levine: Maintaining a Reputation when Strategies are Imperfectly Observed, Review of Economic Studies, **59** (1992), 561–579



Impermanent Reputations

Reputation Effect

Perfect monitoring:

- Ω is the set of all possible outcomes $\omega = \left(\vartheta, a^0, a^1, \ldots\right)$.
- Player 2 makes no mistakes in the inference on player 1's type.
- Beliefs $\mu(h^t; \vartheta_{\widehat{a}_1})$ are non-decreasing for outcomes in

$$\Omega' = \left\{\omega \in \Omega \;\middle|\; A_1^t(\omega) = \widehat{a}_1 \text{ for all } t\right\}.$$

- We used the correlation between player 1's type and the chosen actions to establish a bound on the frequency, with which $\widehat{a}_1$ is expected.

Imperfect monitoring:

- Beliefs fluctuate because bad signals occur even for commitment types.
- What can we say about the evolution of player 2's beliefs?
- The results need some additional concepts from probability theory.

Probability Space

Definition 2.2'

A probability space $(\Omega, \mathcal{F}, P)$ consists of

1. A set Ω of states of the world.
2. A σ-algebra $\mathcal{F}$ of observable events $Y \subseteq \Omega$.
3. A σ-additive measure P with $P(\Omega) = 1$, called probability measure.

Probability space in our setting:

- An outcome under imperfect monitoring is $\omega = (\vartheta, (a^0, y^0), (a^1, y^1), \ldots)$.
- Typically, Ω is the set of all possible outcomes, $\mathcal{F}$ is the σ-algebra of cylinder sets,[1] and P is any probability measure on $\mathcal{F}$ with marginal μ_0 on Θ and conditional distribution of $Y^t \mid A^t = a$ given by $\pi(\cdot \mid a)$.
- P is the probability measure of exogenous events.

[1] A cylinder set is the finite intersection of events $\{A_i^t \in B_i^t\}$, $\{Y^t \in C^t\}$, and $\{\theta \in D\}$ for measurable $B_i^t \subseteq \mathcal{A}_i$, $C^t \subseteq \mathcal{Y}$, and $D \subseteq \Theta$.

Player 2's Information

Player 2's information:

- Player 2 observes $(H_2^t)_{t\geq 0}$, where $H_2^t = (A_2^0, Y^0, \dots, A_2^{t-1}, Y^{t-1})$.
- At time t, given history h_2^t, player 2 knows that the outcome ω is an outcome, in which $H_2^t(\omega) = h_2^t$.

Finite case:

- Suppose that Θ, $\mathcal{A}$, and $\mathcal{Y}$ are finite.
- Before time t has been reached, player 2 does not know $H_2^t(\omega)$ yet.
- Player 2 knows, however, that his/her information will be a partition of Ω into sets of the form $\{\omega \in \Omega \mid H_2^t(\omega) = h_2^t\}$.
- The σ-algebra generated by H_2^t, denoted by $\mathcal{F}_2^t$, contains all events $\{\omega \in \Omega \mid H_2^t(\omega) = h_2^t\}$ and any unions of these events.

Flow of Information

General case:

- Player 2 knows that his/her information partitions Ω into cylinder sets

$$\left\{A_2^0 \in B_2^0\right\} \cap \left\{Y^0 \in C^0\right\} \cap \ldots \cap \left\{A_2^{t-1} \in B_2^{t-1}\right\} \cap \left\{Y^{t-1} \in C^{t-1}\right\}$$

  where $B_2^s \subseteq \mathcal{A}_2$ and $C^s \subseteq \mathcal{Y}$ are (typically Borel-) measurable sets.
- The σ-algebra generated by H_2^t, again denoted by $\mathcal{F}_t^2$, contains any countable unions of these cylinder sets.

Flow of information:

- Information flow is captured by a filtration $\mathbb{F}_2 = (\mathcal{F}_2^t)_{t\geq 0}$, where $\mathcal{F}_2^t$ is the σ-algebra generated by H_2^t.
- Since partitions get finer as time progresses, we have $\mathcal{F}_2^t \subseteq \mathcal{F}_2^{t+1}$, showing that information accumulates over time.

Conditional Expectation

Conditional expectation with respect to σ-algebra:

- For any random variable X, the conditional expectation $\mathbb{E}[X \mid \mathcal{F}_2^t]$ is an $\mathcal{F}_2^t$-measurable random variable $\widehat{X}$ such that for every $Z \in \mathcal{F}_2^t$,
$$\mathbb{E}[\widehat{X} 1_Z] = \mathbb{E}[X 1_Z].$$
- Tower property: for any random variable X and any $s \leq t$, we have
$$\mathbb{E}[X \mid \mathcal{F}_2^s] = \mathbb{E}\left[\mathbb{E}\left[X \,\middle|\, \mathcal{F}_2^t\right] \,\middle|\, \mathcal{F}_2^s\right].$$

Relation to conditioning on histories:

- $\mathbb{E}[X \mid \mathcal{F}_2^t]$ is a random variable, whereas $\mathbb{E}[X \mid H_2^t = h_2^t]$ is a constant.
- $\mathbb{E}[X \mid \mathcal{F}_2^t]$ takes value $\mathbb{E}[X \mid H_2^t = h_2^t]$ on the set $\{H_2^t = h_2^t\}$.
- You can switching between the two via
$$\mathbb{E}\left[X \,\middle|\, \mathcal{F}_2^t\right] = \mathbb{E}\left[X \,\middle|\, H_2^t = h_2^t\right]\Big|_{h_2^t = H_2^t} = \sum_{h_2^t \in \mathcal{H}_2^t} \mathbb{E}\left[X \,\middle|\, H_2^t = h_2^t\right] 1_{\{H_2^t = h_2^t\}}.$$

Belief Process

Naive way to define beliefs:

- Define $\mu(h_2^t) \in \Delta(\Theta)$ by setting $\mu(h_2^t; \vartheta) = P_\sigma(\theta = \vartheta \,|\, H_2^t = h_2^t)$.
- Define the belief process $\mu = (\mu_t)_{t \geq 0}$ by setting $\mu_t = \mu(H_2^t)$.
- This is well-defined only for histories h_2^t with positive measure.

Belief processes:

- Define belief process as $\Delta(\Theta)$-valued stochastic sequence $\mu = (\mu_t)_{t \geq 0}$ by setting $\mu_t(\vartheta) := P_\sigma(\theta = \vartheta \,|\, \mathcal{F}_2^t)$ for every $\vartheta \in \Theta$.
- For a history h_2^t with positive measure, we have

$$\begin{aligned}\mu_t 1_{\{H_2^t = h_2^t\}} &= P_\sigma\big(\theta = \vartheta \,\big|\, \mathcal{F}_2^t\big) 1_{\{H_2^t = h_2^t\}} \\ &= P_\sigma\big(\theta = \vartheta \,\big|\, H_2^t = h_2^t\big) 1_{\{H_2^t = h_2^t\}} = \mu(h_2^t) 1_{\{H_2^t = h_2^t\}}.\end{aligned}$$

Martingales

Definition 12.2

A sequence of random variables $X = (X_t)_{t \geq 0}$ is a martingale if

1. X is adapted to $\mathbb{F}_2$, that is, X_t is $\mathcal{F}_2^t$-measurable for every $t \geq 0$,
2. X is integrable, i.e., $\mathbb{E}[|X_t|] < \infty$ for every $t \geq 0$,
3. X satisfies the martingale property:

$$\mathbb{E}[X_t \,|\, \mathcal{F}_2^s] = X_s \qquad \text{for all } s \leq t.$$

Remark:

- The martingale property means that X is constant on average.
- If X satisfies properties 1. and 2. but property 3. holds with $\leq$ ($\geq$), then X is called a supermartingale (submartingale), respectively.
- A supermartingale is decreasing, a submartingale increasing on average.

Martingale Properties of Belief Process

Belief process:

- Let $\widehat{\alpha}_1$ be the commitment action player 1 wishes to mimick with commitment type $\vartheta_{\widehat{\alpha}_1} \in \Theta_c$ with positive prior $\widehat{\mu}_0 = \mu\big(\vartheta_{\widehat{\alpha}_1}\big) > 0$.
- Define player 2's belief process $(\widehat{\mu}_t)_{t\geq 0}$ that player 1 is commitment type $\vartheta_{\widehat{\alpha}_1}$ by setting $\widehat{\mu}_t := P_\sigma(\theta = \theta_{\widehat{\alpha}_1} \,|\, \mathcal{F}_2^t)$.
- It follows immediately from the tower property of conditional expectations that $\widehat{\mu}$ is a $(P_\sigma, \mathbb{F}_2)$-martingale.

Belief process under deviation:

- For any strategy profile σ, let

$$\widehat{P}_\sigma = P_\sigma(\,\cdot\,|\,\theta = \vartheta_{\widehat{\alpha}_1}), \qquad \widetilde{P}_\sigma = P_\sigma(\,\cdot\,|\,\theta \neq \vartheta_{\widehat{\alpha}_1}).$$

- What properties does $\widehat{\mu}$ have under $\widehat{P}_\sigma$ or $\widetilde{P}_\sigma$?

Martingale Properties of Belief Process

Lemma 12.3

For any strategy profile σ, the belief process $(\widehat{\mu}_t)_{t\geq 0}$ is a P_σ-martingale, a $\widehat{P}_\sigma$-submartingale, and a $\widetilde{P}_\sigma$-supermartingale.

Interpretation:

- Beliefs are constant on average under P_σ.
- Beliefs increase on average if player 1 is the commitment type.
- Beliefs decrease on average if player 1 is not the commitment type.

Proof, step 1: $(\widehat{\mu}_t)_{t\geq 0}$ is a P_σ-martingale:

$$\mathbb{E}_\sigma\left[\widehat{\mu}_{t+1} \,\middle|\, \mathcal{F}_2^t\right] = \mathbb{E}_\sigma\left[\mathbb{E}_\sigma\left[1_{\{\theta=\theta_{\widehat{\alpha}_1}\}} \,\middle|\, \mathcal{F}_2^{t+1}\right] \,\middle|\, \mathcal{F}_2^t\right] = \mathbb{E}_\sigma\left[1_{\{\theta=\theta_{\widehat{\alpha}_1}\}} \,\middle|\, \mathcal{F}_2^t\right] = \widehat{\mu}_t.$$

Change of Probability Measures

Lemma 12.4

Let Z be a random variable on a probability space $(\Omega, \mathcal{F}, P)$ with $\mathbb{E}[Z] = 1$. Then $Q(B) := \mathbb{E}[Z 1_B]$ for $B \in \mathcal{F}$ defines a new probability measure with

$$\mathbb{E}_Q\left[X \,\middle|\, \mathcal{F}_2^t\right] = \frac{\mathbb{E}[XZ \mid \mathcal{F}_2^t]}{\mathbb{E}[Z \mid \mathcal{F}_2^t]}.$$

The process $(Z_t)_{t \geq 0}$ defined by $Z_t := \mathbb{E}[Z \mid \mathcal{F}_2^t]$ is called the density process and we write $Z = \frac{\mathrm{d}Q}{\mathrm{d}P}$ and $Z_t = \frac{\mathrm{d}Q_t}{\mathrm{d}P_t}$.

- For $\widehat{P}_\sigma = P_\sigma(\,\cdot \mid \theta = \vartheta_{\hat{a}_1})$, the density process is

$$\frac{\mathrm{d}\widehat{P}_\sigma}{\mathrm{d}P_\sigma} = \frac{1_{\{\theta = \vartheta_{\hat{a}_1}\}}}{P_\sigma(\theta = \vartheta_{\hat{a}_1})} = \frac{1_{\{\theta = \vartheta_{\hat{a}_1}\}}}{\widehat{\mu}_0}, \qquad \frac{\mathrm{d}\widehat{P}_{\sigma,t}}{\mathrm{d}P_{\sigma,t}} = \frac{\widehat{\mu}_t}{\widehat{\mu}_0}.$$

- For $\widetilde{P}_\sigma = P_\sigma(\,\cdot \mid \theta \neq \vartheta_{\hat{a}_1})$ we get $\frac{\mathrm{d}\widetilde{P}_\sigma}{\mathrm{d}P_\sigma} = \frac{1_{\{\theta \neq \vartheta_{\hat{a}_1}\}}}{1 - \widehat{\mu}_0}$ and $\frac{\mathrm{d}\widetilde{P}_{\sigma,t}}{\mathrm{d}P_{\sigma,t}} = \frac{1 - \widehat{\mu}_t}{1 - \widehat{\mu}_0}$

Proof

Step 2: The process $\left(\frac{\widehat{\mu}_t}{1-\widehat{\mu}_t}\right)_{t\geq 0}$ is a $\widetilde{P}_\sigma$-martingale.

$$\widetilde{\mathbb{E}}_\sigma\left[\frac{\widehat{\mu}_{t+1}}{1-\widehat{\mu}_{t+1}}\,\middle|\,\mathcal{F}_2^t\right] = \frac{\mathbb{E}_\sigma\left[\frac{\mathrm{d}\widetilde{P}_\sigma}{\mathrm{d}P_\sigma}\frac{\widehat{\mu}_{t+1}}{1-\widehat{\mu}_{t+1}}\,\middle|\,\mathcal{F}_2^t\right]}{\mathbb{E}_\sigma\left[\frac{\mathrm{d}\widetilde{P}_\sigma}{\mathrm{d}P_\sigma}\,\middle|\,\mathcal{F}_2^t\right]} = \frac{\mathbb{E}_\sigma\left[1_{\left\{\theta\neq\vartheta_{\widehat{a}_1}\right\}}\frac{\widehat{\mu}_{t+1}}{1-\widehat{\mu}_{t+1}}\,\middle|\,\mathcal{F}_2^t\right]}{1-\widehat{\mu}_t}$$

$$= \frac{\mathbb{E}_\sigma\left[\mathbb{E}_\sigma\left[1_{\left\{\theta\neq\vartheta_{\widehat{a}_1}\right\}}\frac{\widehat{\mu}_{t+1}}{1-\widehat{\mu}_{t+1}}\,\middle|\,\mathcal{F}_{t+1}^2\right]\,\middle|\,\mathcal{F}_2^t\right]}{1-\widehat{\mu}_t}$$

$$= \frac{\mathbb{E}_\sigma\left[\frac{\widehat{\mu}_{t+1}}{1-\widehat{\mu}_{t+1}}\mathbb{E}_\sigma\left[1_{\left\{\theta\neq\vartheta_{\widehat{a}_1}\right\}}\,\middle|\,\mathcal{F}_{t+1}^2\right]\,\middle|\,\mathcal{F}_2^t\right]}{1-\widehat{\mu}_t}$$

$$= \frac{\mathbb{E}_\sigma\left[\widehat{\mu}_{t+1}\,\middle|\,\mathcal{F}_2^t\right]}{1-\widehat{\mu}_t} = \frac{\widehat{\mu}_t}{1-\widehat{\mu}_t}.$$

Proof

Step 3: $(\widehat{\mu}_t)_{t\geq 0}$ is a $\widetilde{P}_\sigma$-supermartingale

- Jensen's inequality says that $\mathbb{E}[\varphi(X)] \geq \varphi(\mathbb{E}[X])$ for convex φ.
- Note that $\varphi\left(\frac{x}{1-x}\right) = -x$ is convex, hence

$$\widetilde{\mathbb{E}}\left[-\widehat{\mu}_{t+1} \,\middle|\, \mathcal{F}_2^t\right] \geq \varphi\left(\widetilde{\mathbb{E}}\left[\frac{\widehat{\mu}_{t+1}}{1-\widehat{\mu}_{t+1}} \,\middle|\, \mathcal{F}_2^t\right]\right) = \varphi\left(\frac{\widehat{\mu}_t}{1-\widehat{\mu}_t}\right) = -\widehat{\mu}_t.$$

- Therefore, $(\widehat{\mu}_t)_{t\geq 0}$ is a $\widetilde{P}_\sigma$-supermartingale.

Step 4: $(\widehat{\mu}_t)_{t\geq 0}$ is a $\widehat{P}_\sigma$-submartingale probability

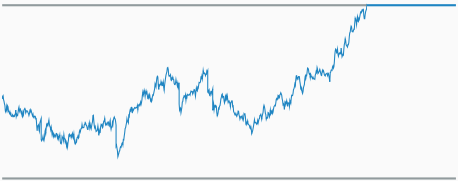
- By the law of total probability,

$$\begin{aligned}\widehat{\mu}_t = P_\sigma\left(\widehat{\mu}_{t+1} \,\middle|\, \mathcal{F}_2^t\right) &= \widehat{\mu}_t \widehat{P}_\sigma\left(\widehat{\mu}_{t+1} \,\middle|\, \mathcal{F}_2^t\right) + (1-\widehat{\mu}_t)\widetilde{P}_\sigma\left(\widehat{\mu}_{t+1} \,\middle|\, \mathcal{F}_2^t\right) \\ &\leq \widehat{\mu}_t \widehat{P}_\sigma\left(\widehat{\mu}_{t+1} \,\middle|\, \mathcal{F}_2^t\right) + (1-\widehat{\mu}_t)\widehat{\mu}_t\end{aligned}$$

Martingale Convergence Theorem

Theorem 12.5 (Martingale Convergence Theorem)

Any bounded martingale converges P-almost-surely.



Intuition:

- Because a martingale is constant in expectation, expected upward movements are equal to the expected downward movements.
- Once the martingale hits the upper boundary, no upward movements are possible anymore, hence there are no downward movements either.

Identifiability of Deviations

Definition 12.6

Suppose that $\mathcal{Y}$ and $\mathcal{A}_i$ for $i = 1, 2$ are finite. We say that π:

1. has full support if $\pi(y \mid a) > 0$ for every $y \in \mathcal{Y}$ and every $a \in \mathcal{A}$,
2. is linearly independent if, for both players i and every action $a_i \in \mathcal{A}_i$, the vectors $\pi(\cdot \mid a_i, a_{-i})$ for $a_{-i} \in \mathcal{A}_{-i}$ are linearly independent.

Remark:

- Linear independence ensures that player i, knowing he/she played a_i, can statistically distinguish $\alpha_{-i} \neq \alpha'_{-i}$ by observing the signal.
- The signal in perfect monitoring games is $Y = A$, hence $\pi(\cdot \mid a_i, a_{-i})$ are orthogonal and, in particular, linearly independent.
- Perfect monitoring does not satisfy the full support assumption.

Impermanent Reputations

Proposition 12.7

Suppose that $\mathcal{Y}$ and $\mathcal{A}_i$ for $i = 1, 2$ are finite, that π has full support, and that π is linearly independent. Let $\widehat{\alpha}_1$ be a commitment action with a unique myopic best reply $\widehat{a}_2$ such that $(\widehat{\alpha}_1, \widehat{a}_2)$ is not a stage-game Nash equilibrium. Then in any Nash equilibrium σ of the reputation game,

$$P_\sigma\left(\theta = \vartheta_{\widehat{\alpha}_1} \,\middle|\, \mathcal{F}_2^t\right) \to 0 \qquad \widetilde{P}_\sigma\text{-a.s.}$$

Interpretation:

- If the signal is sufficiently informative and player 1 is not the commitment type, then a reputation effects are temporary in any equilibrium.
- This may be surprising since identifiability helps building a reputation.

Idea of Proof

Idea of proof:

- Fix a Nash equilibrium σ. Suppose that there exists a set of outcomes $\Omega'' \subseteq \Omega$ with $\widetilde{P}(\Omega'') = P_\sigma(\Omega'' \,|\, \theta \neq \vartheta_{\widehat{\alpha}_1}) > 0$ such that on Ω'':

$$\lim_{t\to\infty} P_\sigma\left(\theta = \vartheta_p \,\middle|\, \mathcal{F}_2^t\right) > 0, \qquad \lim_{t\to\infty} P_\sigma\left(\theta = \vartheta_{\widehat{\alpha}^1} \,\middle|\, \mathcal{F}_2^t\right) > 0.$$

- On Ω'', player 2 cannot distinguish signal distributions and must believe both types are playing the same strategy on average.
- Linear independence: player 2 will eventually best reply to $\widehat{\alpha}_1$ and player 1 will eventually learn that player 2 is best replying to $\widehat{\alpha}_1$.
- Since $(\widehat{\alpha}_1, \widehat{a}_2)$ is not a Nash equilibrium and beliefs are constant in the limit, player 1 must find it profitable to deviate eventually.
- This contradicts that player 2's beliefs are constant in the limit.

Summary

Imperfect monitoring:

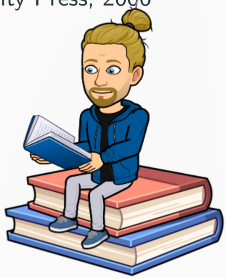
- Building a reputation may take longer than with perfect monitoring.
- Once the reputation is built and player 1 learns that the reputation is effective, he/she exploits the reputation to reap its benefits.
- This causes reputation effects to be temporary.
- We say that reputations are optimally depleted.

Relaxing the full-support assumption:

- Let $\mathcal{Y}(\alpha)$ denote the signals y with $\pi(y \mid \alpha) > 0$.
- The result still holds if player 1 has a profitable deviation $\widetilde{\alpha}_1$ to $(\widehat{\alpha}_1, \widehat{a}_2)$ such that $\mathcal{Y}(\widetilde{\alpha}_1, \widehat{a}_2) \subseteq \mathcal{Y}(\widehat{\alpha}_1, \widehat{a}_2)$.

Literature

- G.J. Mailath and L. Samuelson: Repeated Games and Reputations: Long-Run Relationships, Chapters 15.5–15.6, Oxford University Press, 2006
- M.W. Cripps, G.J. Mailath, and L. Samuelson: Imperfect Monitoring and Impermanent Reputations, Econometrica, **72** (2004), 407–432



Reputation Bound with Imperfect Monitoring

Reputation Bound

Reputation bound with perfect monitoring:

- If the payoff type incessantly plays a commitment action $\widehat{a}_1$, player 2 must eventually best reply to $\widehat{a}_1$ by the reputation effect.
- Any Nash equilibrium σ must be robust to the deviation $\widetilde{\sigma}_1$ that mimics $\vartheta_{\widehat{a}_1}$, hence the payoff induced by $\widetilde{\sigma}_1$ is a lower bound for $U_1(\sigma)$.
- Note that this proof does not require that $\widetilde{\sigma}_1$ is a Nash equilibrium.

Reputation effect:

- For player 2 to eventually best respond to $\widehat{a}_1$, we need that:
  - Player 2 can distinguish $\widehat{a}_1$ from other actions,
  - $\mathcal{A}_2$ is finite, so that when player 2 expects to see $\widehat{a}_1$ with probability above some threshold ζ, he/she will play a best response to $\widehat{a}_1$.
- Linear independence should allow us to recover the reputation effect.

Statistical Indistinguishability

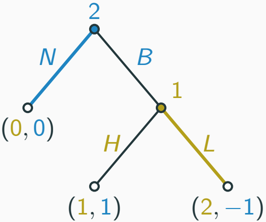
Imperfect monitoring:

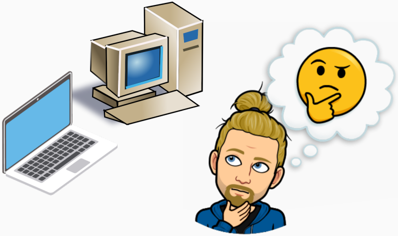
- In a Nash equilibrium, reputations are impermanent because player 1 will eventually find it profitable to deviate.
- Since the reputation bound does not rely on a Nash equilibrium, but the deviation to $\widetilde{\sigma}_1$, maybe weaker requirements on π are sufficient.

Statistical indistinguishability:

- If (α_1, α_2) and $(\widetilde{\alpha}_1, \alpha_2)$ lead to the same signal distribution, player 2 is not able to differentiate between repeated play of α_1 and $\widetilde{\alpha}_1$.
- Payoff type may imitate ϑ_{α_1} but get best responses to $\widetilde{\alpha}_1$.
- This can be very unfortunate for player 1.

Sequential Product Choice





Dynamic stage game:

- Customers choose whether to (B)uy or (N)ot to buy the product.
- If the Customers choose to buy, the Firm decides whether to exert (H)igh or (L)ow effort in producing/delivering the product.
- The terminal node of the stage game is observed: $Y \in \{y_N, y_{BH}, y_{BL}\}$.
- Monitoring is imperfect because no information about the Firm is revealed if Customers choose N.

Signal Distribution in Sequential Product Choice

	H, B	H, N	L, B	L, N
y_N	0	1	0	1
y_{BH}	1	0	0	0
y_{BL}	0	0	1	0

Repeated play of H:

- If the Customers never buy the product, they do not observe whether the Firm chooses H even if $\sigma_1(\vartheta_p, h) = H$ for every history h.
- Since H and L are indistinguishable on the path, the Customers may best reply to what they perceive as L by playing N forever.
- The Firm cannot exploit the reputation effect.
- If $\mu(\vartheta_H) < \frac{1}{2}$, then $(0, 0)$ is an equilibrium payoff even if $\delta > \frac{1}{2}$.

Epsilon-Confirmed Best Responses

Definition 12.8

For any $\varepsilon > 0$, an action $\alpha_2 \in \Delta(\mathcal{A}_2)$ is an ε-confirmed best response to $\alpha_1 \in \Delta(\mathcal{A}_1)$ if there exists $\widetilde{\alpha}_1 \in \Delta(\mathcal{A}_1)$ such that:

1. $\alpha_2 \in \mathcal{B}_2(\widetilde{\alpha}_1)$.
2. $\|\pi(\,\cdot \mid \alpha_1, \alpha_2) - \pi(\,\cdot \mid \widetilde{\alpha}_1, \alpha_2)\| \leq \varepsilon$.

Let $\mathcal{B}_2^{\varepsilon}(\alpha_1)$ denote the set of all ε-confirmed best responses to α_1.

Remark:

- This notion measures closeness in signal distribution (hence ε-confirmed best responses), not closeness in utility.
- If strategic player repeatedly plays α_1, for any $\varepsilon > 0$, he/she can expect responses from the set $\mathcal{B}_2^{\varepsilon}(\alpha_1)$ in the long run.

Signal Distribution in Sequential Product Choice

	H, B	H, N	L, B	L, N
y_N	0	1	0	1
y_{BH}	1	0	0	0
y_{BL}	0	0	1	0

0-confirmed best response:

- (H, N) and (L, N) lead to the same signal with probability 1, hence

$$\left\|\pi(\cdot \mid H, N) - \pi(\cdot \mid L, N)\right\| = 0.$$

- N is a best response to L, hence a 0-confirmed best response to H.

Linear Independence

Lemma 12.9

If π is linearly independent, then $\mathcal{B}_2^0(\alpha_1) = \mathcal{B}_2(\alpha_1)$ for any α_1, that is, 0-confirmed best responses are best responses. Moreover, if $\mathcal{Y}$ and $\mathcal{A}_i$ for $i = 1, 2$ are finite, then $\mathcal{B}_2^{\varepsilon}(\alpha_1) = \mathcal{B}_2(\alpha_1)$ for sufficiently small ε.

Proof:

- By linear independence, $\pi(\cdot \mid \alpha_2, \alpha_1)$ is injective in α_1 for any α_2.
- If $\mathcal{A}_2$ is finite, then for any α_1, there exists $\varepsilon(\alpha_1) > 0$ such that $\mathcal{B}_2(\widetilde{\alpha}_1) \subseteq \mathcal{B}_2(\alpha_1)$ for any $\widetilde{\alpha}_1$ with $\|\widetilde{\alpha}_1 - \alpha_1\| \leq \varepsilon(\alpha_1)$.
- Since $\pi(\cdot \mid \alpha_2, \alpha_1)$ is continuous in α_1, there exists $\varepsilon > 0$ such that

$$\|\pi(\cdot \mid \alpha_1, \alpha_2) - \pi(\cdot \mid \widetilde{\alpha}_1, \alpha_2)\| \leq \varepsilon$$

  implies $\|\widetilde{\alpha}_1 - \alpha_1\| \leq \varepsilon(\alpha_1)$, hence $\mathcal{B}_2(\widetilde{\alpha}_1) \subseteq \mathcal{B}_2(\alpha_1)$.

Reputation Bound with Imperfect Monitoring

Theorem 12.10

Suppose that $\mu_0(\vartheta_p) > 0$. For any $\varepsilon > 0$ and any $\widehat{\alpha}_1 \in \Delta(\mathcal{A}_1)$ with $\vartheta_{\widehat{\alpha}_1} \in \Theta_c$ and $\mu_0(\vartheta_{\widehat{\alpha}_1}) > 0$, there exists a constant $K(\mu_0) > 0$ such that

$$\inf_{Eq.\ \sigma} U_1(\sigma) \geq (1-\varepsilon)\delta^K \min_{\alpha_2 \in \mathcal{B}_2^\varepsilon(\widehat{\alpha}_1)} u_1(\widehat{\alpha}_1, \alpha_2) + (1-(1-\varepsilon)\delta^K) \min_{a \in \mathcal{A}} u_1(a).$$

Interpretation:

- It takes $K(\mu_0)$ periods to build a sufficient reputation to get ε-confirmed best responses to the commitment action $\widehat{\alpha}_1$.
- If π is linearly independent, the long-lived player elicit best responses to $\widehat{\alpha}_1$ instead of ε-confirmed best responses.
- The proof is similar to Theorem 11.8 and can be found in the book.

University Reputation

UNIVERSITÉ DU LUXEMBOURG

University of Luxembourg:

- If they are the strategic type, they will not fool us asymptotically.
- But: deviating will be in the university's own interest.
- Until the university of Luxembourg finds it beneficial to reveal its type, students actually do get a good education.

Check Your Understanding

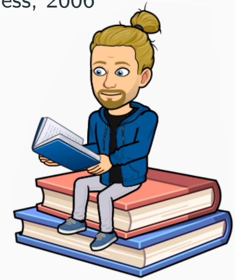
True or false:

1. Reputations are impermanent because they are too hard to build with imperfect monitoring.
2. Suppose commitment type $\vartheta_{\widehat{a}_1}$ maximizes the long-lived player's payoff in the long-run. Other commitment types do not have an impact on payoffs.
3. When attaining the reputation bound in Theorem 12.10, the long-lived player earns a low payoff during the first $K(\mu_0)$ periods.
4. Suppose all actions are observed. To justify a lower bound of $\underline{v}(\widehat{\alpha}_1)$ on equilibrium payoffs for a mixed commitment type $\vartheta_{\widehat{\alpha}_1}$ we need the imperfect-monitoring reputation bound.

Scan me

Literature

- G.J. Mailath and L. Samuelson: Repeated Games and Reputations: Long-Run Relationships, Chapter 15.4, Oxford University Press, 2006
- D. Fudenberg and D.K. Levine: Maintaining a Reputation when Strategies are Imperfectly Observed, Review of Economic Studies, **59** (1992), 561–579



Reputations with Long-Lived Players

Motivation for Reputation Games

Short-lived players:

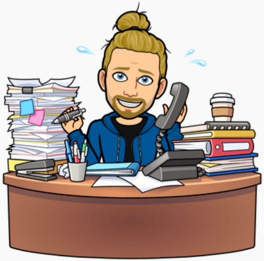
- Can a sequence of customers, citizens, etc. come to expect certain behavior because they have observed it repeatedly?
- How can we connect the continuation game in a meaningful way to players' observation without sacrificing tractability?

Original motivation:

- Finitely repeated games with a unique stage-game Nash equilibrium have a unique SPE, which is often deemed unintuitive.
- If players believed that there was a small chance of facing a cooperative type, would an outcome be observed that is closer to our expectation?

Partnership Game

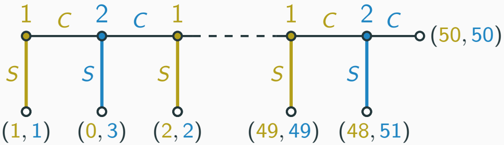
	E	S
E	$2, 2$	$1, 3$
S	$3, 1$	$0, 0$



Group project:

- With complete information, the unique SPE predicts that there will be no effort, even if the horizon $T < \infty$ is potentially very long.
- Intuitively, we expect behavior closer to grim trigger if T is large.
- Does the set of equilibria change if players believe the other is a hard-working "grim-trigger type" with some small prior probability?

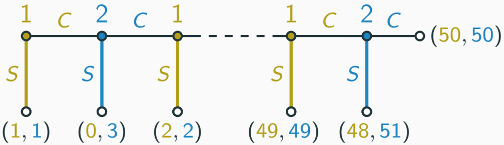
Failure to Replicate Subgame Perfection in the Lab



Centipede game:

- Suppose you are in Joseph's lab, participating in an experiment whether individuals act in a subgame perfect way.
- A typical such lab experiment is the centipede game, in which players iteratively choose to either (C)ontinue or (S)top the game.
- Lab experiments typically do not confirm the unique SPE.
- One possible explanation is that we believe there are commitment types who choose to continue at least once.

Modeling Strategic Uncertainty



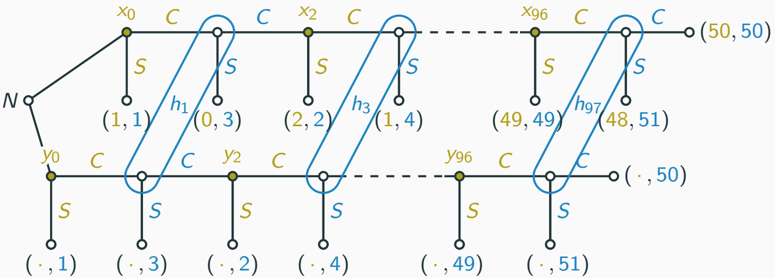
Commitment types:

- Consider a set of commitment types ϑ_t for each $t \geq 0$, who choose C at their first t decision nodes and S forever after.
- If we place sufficiently high beliefs on facing a commitment type, choosing C in the first period becomes uniquely optimal.

Reputations:

- For what priors will a reputation be built in this game?
- How long is the reputation maintained? $\rightarrow$ Let's find out!

Setting up the Reputation Game



Commitment types:

- Suppose that Player 1 is either a payoff type or a commitment type ϑ_C who chooses C at every node.
- Player 2 has only one type and places prior $\mu_0 > 0$ on facing ϑ_C.
- For simplicity of notation, write $u_1(x_n, S)$ and $u_2(h_n, S)$ the utility for stopping at node x_n or information set h_n, respectively.

Strict Improvement Every Two Periods

Claim 1

Consider any perfect Bayesian equilibrium $\sigma = (\sigma_1, \sigma_2)$.

1. *If $\sigma_2(h_n) = C$, then $\sigma_1(\vartheta_p, x_{n-1}) = C$.*
2. *If $\sigma_2(h_n) = S$, then $\sigma_1(\vartheta_p, x_{n-1}) = S$.*

Proof of statement 1:

- If $\sigma_2(h_n) = C$, then $\sigma_1(x_{n-1}) = C$ yields a utility of at least

$$u_1(\sigma) \geq u_1(x_{n+1}, S) > u_1(x_{n-1}, S).$$

Proof of statement 2:

- If $\sigma_2(h_n) = S$, then choosing $\sigma_1(x_{n-1}) = C$ yields a utility of

$$u_1(x_{n-1}, S) - 1 < u_1(x_{n-1}, S).$$

Strict Improvement Every Two Periods

Claim 2

If $\sigma_2(h_n) = C$ for some n in a PBE σ, then $\sigma_2(h_k) = C$ and $\sigma_1(\vartheta_p, x_k) = C$ for $k < n$. In particular, $\mu(h_n) = \mu_0$.

Proof by induction:

- If $\sigma_2(h_n) = C$ for some n, then both types of Player 1 choose C at x_{n-1} or y_{n-1}, respectively, by Claim 1.
- Since σ is a PBE, choosing $\sigma_2(h_{n-2}) = C$ yield a utility of at least

$$u_2(\sigma) \geq u_2(h_n, S) > u_2(h_{n-2}, S).$$

- By backward induction, $\sigma_2(h_k) = C$ for all $k < n$.
- By Claim 1, $\sigma_1(\vartheta_p, x_k) = C$ for all $k < n$.

Climbing the Centipede

Claim 3

In any perfect Bayesian equilibrium σ, we have $\sigma_2(h_n; C) > 0$ except for $n = 97$, where $\sigma_2(h_{97}) = S$.

Proof:

- Suppose towards a contradiction that $\sigma_2(h_n; C) = 0$ for $n < 97$.
- Then $\sigma_1(\vartheta_p, x_{n-1}) = S$ by Claim 1.
- Thus, conditional on reaching information set h_n, Player 2 must place beliefs 1 on facing the commitment type.
- Therefore, continuing at h_n is optimal with the intent to stop at h_{97}.

Climbing the Centipede

Claim 4

In any perfect Bayesian equilibrium σ, we have $\sigma_1(\vartheta_p, x_n; C) > 0$ except for $n = 96$, where $\sigma_1(\vartheta_p, x_{96}) = S$.

Proof:

- Suppose towards a contradiction that $\sigma_1(\vartheta_p, x_n; C) = 0$ for $n < 96$.
- Then, conditional on reaching information set h_{n+1}, Player 2 must place beliefs 1 on facing the commitment type.
- Therefore, continuing at h_n is optimal with the intent to stop at h_{97}.
- Then $\sigma_1(\vartheta_p, x_n) = C$ by Claim 1.

Three Equilibrium Phases

Indifference:

- At every information set h_n, in which Player 2 mixes, he/she must receive a utility of $u_2(\sigma) = u_2(h_n, S)$ by the indifference principle.
- Equating $u_2(h_n, S)$ to the expected utility of continuing yields

$$u_2(h_n, S) = \big(\mu(h_n) + \sigma_1(\vartheta_p, x_{n+1}; C)(1 - \mu(h_n))\big) u_2(h_{n+2}, S)$$

$$+ \sigma_1(\vartheta_p, x_{n+1}; S)(1 - \mu(h_n))\big(u_2(h_n, S) - 1\big).$$

- Using $u_2(h_{n+2}, S) = u_2(h_n, S) + 1$, we can solve this for

$$\sigma_1(\vartheta_p, x_{n+1}; S) = \frac{1}{2(1 - \mu(h_n))}.$$

Three Equilibrium Phases

Indifference:

- The indifference principle for Player 1 yields

$$u_1(x_n, S) = \sigma_2(h_{n+1}; C)u_1(x_{n+2}, S) + \sigma_2(h_{n+1}; S)\big(u_1(x_n, S) - 1\big),$$

which we solve for $\sigma_2(h_{n+1}; S) = \frac{1}{2}$.

Beliefs:

- Bayesian updating implies that

$$\mu(h_{n+2}) = \frac{\mu(h_n)}{\mu(h_n) + (1 - \sigma_1(\vartheta_p, x_{n+1}; S))(1 - \mu(h_n))} = \frac{\mu(h_n)}{2}.$$

- Since $\mu(h_{97}) = 1$, we obtain $\mu(h_n) = \left(\frac{1}{2}\right)^{\frac{97-n}{2}}$ during the mixing phase.

Three Equilibrium Phases

Lemma 12.11

Consider the depicted version of the centipede game with $\mu_0(\vartheta_C) \geq \left(\frac{1}{2}\right)^{48}$. Then the unique perfect Bayesian equilibrium is characterized by

1. *Both players choosing C at x_n, h_n for $n < 97 + \frac{\log(\mu_0)}{\log(2)} =: n_0$.*
2. *Both players mixing at x_n, h_n for $n_0 \leq n < 96$ with probability*

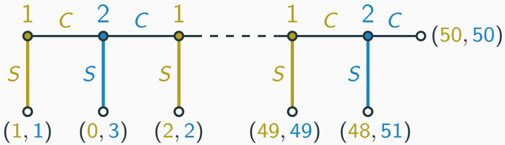
$$\sigma_1(\vartheta_p, x_n; S) = \frac{1}{2 - \left(\frac{1}{2}\right)^{\frac{96-n}{2}}}, \qquad \sigma_2(h_n; S) = \frac{1}{2}.$$

3. *Both players choosing S at x_{96} and h_{97}, respectively.*

Note:

- It takes $K(\mu_0) = -\frac{\log(\mu_0)}{\log(2)}$ periods to build the reputation.
- Observe that $K(\mu_0)$ depends only on μ_0, not on T.

Reputations and Expected Play



Reputations:

- "Reputations" are built at the end of the game.
- Long-lived players understand that, in equilibrium, they will be built. It is thus not necessary to build them in the beginning.

Expected play:

- At the beginning, both players expect to see C with certainty.
- As the game progresses, it becomes more likely that players stop.
- This is precisely what we expect to see in the lab!

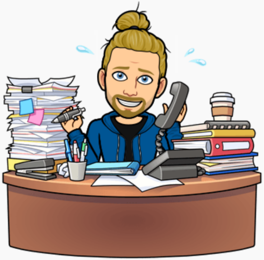
Partnership Game

Proposition 12.12

Suppose that players assign beliefs $\mu_0 > 0$ that their opponent is the grim trigger commitment type in a T-repeated partnership game. Then there exists a constant $K(\mu_0)$, independent of T, such that in any PBE, both players exert effort with certainty in all but the last $K(\mu_0)$ periods.

Remarks:

- For large T, players exert effort in almost every period.
- The proof works along the same lines.
- Because ϑ_G is unforgiving, shirking is equivalent to stopping in thecentipede.



Summary

Finitely repeated games:

- Players cooperate with decreasing probability as the game progresses.
- As a consequence, beliefs of facing the commitment type rise only at the end, i.e., reputations are built at the end.
- Nevertheless, expected play models observed behavior quite well.

Infinitely repeated games:

- Building of beliefs cannot be deferred to the end.
- You will show in the assignment that when there is a single commitment type and perfect monitoring, reputations are built in one period.

Literature

- G.J. Mailath and L. Samuelson: Repeated Games and Reputations: Long-Run Relationships, Chapter 17, Oxford University Press, 2006
- S. Tadelis: Game Theory: An Introduction, Chapter 17, Princeton University Press, 2013
- P. Milgrom, and J. Roberts: Predation, Reputation and Entry Deterrence, Journal of Economic Theory, **27** (1982), 280–312
- D. Kreps, and R. Wilson: Reputation and Imperfect Information, Journal of Economic Theory, **27** (1982), 253–279