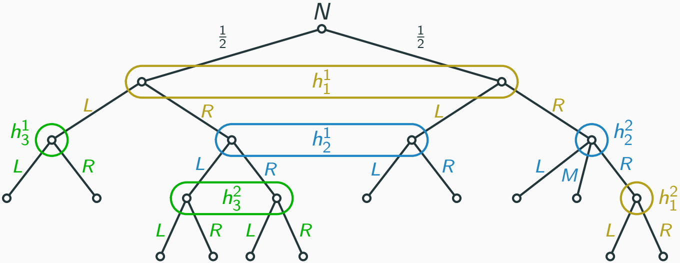
11. Repeated Games and Reputations

ECON 7219 – Games With Incomplete Information

Benjamin Bernard

Extensive-Form Games



Components:

- Assignment from nodes to active player.
- Partition of each player's nodes into information sets h_i.
- Available actions $\mathcal{A}_i(h_i)$ depend on h_i and, for every $x \in h_i$, there is a bijection to successor nodes of x that indicates continuation of play.
- Payoffs depend on terminal payoffs.

Advantages and Disadvantages of Extensive-Form Games

Advantages of extensive-form games:

- It is a very general model, allowing imperfect information due to strategic uncertainty, payoff uncertainty, or imperfect observation.
- A game tree provides a good visual for small games.

Disadvantages of extensive-form games:

- The notion of a game tree inherently restricts the game to be finite.
- With infinite action sets, subgame-perfect equilibria may not exist in very intuitive models like the English auction.
- The notation is quite cumbersome: keeping track of the transition between nodes, information sets, the identity of the active player, etc.
- The generality makes it hard to characterize equilibria analytically.
- If we aim to fit the model to data, it is prone to overspecification.

Beliefs in Extensive-Form Games

Off-path beliefs:

- Off-path beliefs are a nuisance already in 2-player, 2-type, 2-action signaling games. It becomes exponentially worse in longer games.

Payoff vs. information:

- Uninformed players face a trade-off between earning a higher payoff and learning more information at every non-terminal information set.

Bargaining games:

- Sequential bargaining with incomplete information about players' outside options is considered intractable for the above reasons.
- Instead, we use bilateral-trade mechanisms and reputational bargaining.

Examples of Dynamic Games

Some examples of interest:

- Industrial organization: competition, entry deterrence, collusion, etc.
- Collaborations: business partnerships, climate agreements, provision of a public good, upholding of social norms, etc.
- Negotiations: bargaining, legislation, international relations, etc.

Commonality:

- The nature of the interaction does not change drastically over time.
- We can simplify the framework to improve the tractability.

Improving Tractability

Impose additional structure:

- The more structure we impose, the sharper our results will get.
- The benchmark is a repeated game, in which players repeatedly play the same simultaneous-move (or sequential two-move) stage game.[1]

Unfaltering types:

- There will be a single strategic type and a variety of commitment types, who invariably play a commonly known strategy.
- Therefore, off-path belief assign probability 1 to the strategic type.

Short-lived players:

- This week, we assume that there is only one uninformed player, who maximizes their payoffs myopically.
- Thus, there is no trade-off of immediate gains vs. value of information.

[1] The latter is, formally, called a stochastic game.

Repeated Games

Tipping in a Restaurant

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

Tipping:

- The Waiter can exert (E)ffort or he/she can (S)hirk.
- The Client may choose to tip (E) or not to tip (S).
- Isolated interaction: S is the strictly dominant action for each player.

Incentives for tipping:

- The Client is unsure whether he/she will go back in the future.
- Possibility for future retaliation by the Waiter may provide incentives.

Repeated Game

Repeating the stage game:

- Players repeat stage game $\mathcal{G} = (\mathcal{I}, \mathcal{A}, u)$ a total of $T \leq \infty$ times.
- Periods are labeled by $t = 0, \ldots, T-1$ for simplicity of notation.

Available histories:

- At the end of each period, players observe the chosen action profiles, that is,the history h^t at time t is of the form $h^t = (a^0, \ldots, a^{t-1})$.
- Let $\mathcal{H}^t$ denote the set of t-period histories and let $\mathcal{H}(T) = \bigcup_{t=0}^{T} \mathcal{H}^t$.
- We often omit the dependence on T if it is clear from context.

Strategies:

- A pure strategy of player i is a map $\sigma_i : \mathcal{H} \to \mathcal{A}_i$.
- A behavior strategy of player i is a map $\sigma_i : \mathcal{H} \to \Delta(\mathcal{A}_i)$.

Distribution over Outcomes

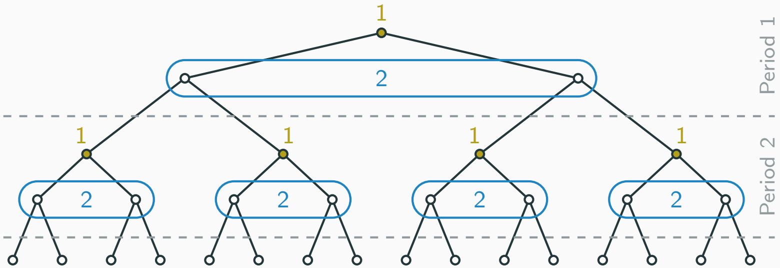
Path of play and outcome:

- The realization of the players' (possibly mixed) actions gives rise to an $\mathcal{A}$-valued stochastic sequence called path of play $A = (A^t)_{t\geq 0}$.
- Conditional on $H^t := (A^0, \ldots, A^{t-1})$, the random variables A_i^t and A_j^t are independent for any $j \neq i$ and any $t \geq 0$.
- $H = (H^t)_{t\geq 0}$ is called the outcome of the game.

Distribution over outcomes:

- Play of strategy profile $\sigma = (\sigma_1, \ldots, \sigma_n)$ induces a probability measure P_σ, under which A_i^t is distributed according to $\sigma_i(H^t)$.
- For finitely repeated games, P_σ is identical to the probability measure we defined for extensive-form games.

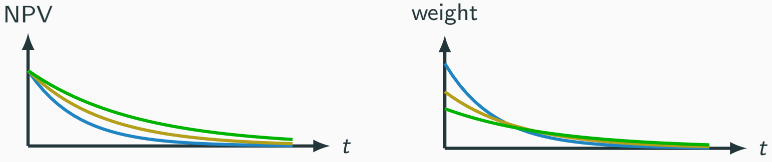
Game Tree



Continuation Games:

- The only proper subgames start in some period t with history h^t.
- The continuation game is identical for each history $h^t \in \mathcal{H}^t$ and every player has the same information h^t in the continuation game.
- Because the game may be infinite, there may be no terminal nodes. Payoffs are earned at the end of each period instead.

Payoffs



Payoffs:

- Players discount future payoffs with a discount factor $\delta \in (0, 1)$.
- Players maximize their average discounted payoff is

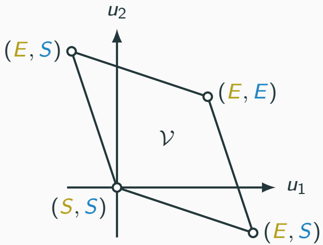
$$U(\sigma) := \frac{1-\delta}{1-\delta^T} \sum_{t=0}^{T-1} \delta^t \mathbb{E}_\sigma \left[u(A^t) \right]. \tag{1}$$

Benefits of considering average payoffs:

- Repeated-game payoffs are in the same units as stage-game payoffs.
- It allows a meaningful comparison between different δ and T.

Feasible Payoffs

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



Feasible payoffs:

- Repeated-game payoffs are a weighted sum of stage-game payoffs.
- Any strategy profile thus attains a payoff vector in $\mathcal{V} := \mathrm{conv}(u(\mathcal{A}))$.
- Conversely, any payoff vector in $\mathcal{V}$ can be attained by some strategy profile in the infinitely repeated game.
- In either case, we call $\mathcal{V}$ the set of feasible payoffs.

How Reasonable Is an Infinitely Repeated Game?

Game with random termination:

- At the end of each period: continue the game with probability δ.
- Note that game ends in finite time with probability 1.
- The players' undiscounted expected payoff in this game is also (1).

Random termination and discounting:

- If, in addition, players discount future payoffs with discount factor ρ, then the discounted expected payoff is

$$U_i(\sigma) = (1 - \delta\rho) \sum_{t=0}^{\infty} (\delta\rho)^t \mathbb{E}_\sigma \left[u_i(A^t) \right].$$

- This is just a repeated game with discount factor $\delta' = \delta\rho$.

Repeated Play of Static Nash Equilibria

Lemma 11.1

Let $\mathcal{A}_{Nash} \subseteq \mathcal{A}$ denote the set of stage-game Nash equilibria, also called static Nash equilibria. Any strategy profile σ with $\sigma(h^t) = \alpha^t \in \mathcal{A}_{Nash}$ for any $h^t \in \mathcal{H}^t$ is a subgame-perfect equilibrium of the repeated game.

Intuition:

- Since α_s is chosen independently of history for any $s \geq t$, nothing any player i does in period t affects i's continuation payoff.
- It is optimal to maximize the period t payoff by playing α_i^{t-1}.

Intertemporal incentives:

- In any subgame-perfect equilibrium with $\sigma(h^t) \notin \mathcal{A}_{Nash}$, players need to be rewarded for playing $\sigma(h^t)$ or punished for not playing $\sigma(h^t)$

Continuation Strategies and Payoffs

Definition 11.2

Consider a behavior strategy profile σ.

1. The continuation strategy $\sigma_i|_{h^t}$ of player i after history h^t is the map $\sigma_i|_{h^t} : \mathcal{H}(T-t) \to \Delta(\mathcal{A}_i)$ with $\sigma_i|_{h^t}(h^s) = \sigma_i(h^t h^s)$.
2. The continuation value after history h^t is

$$U(\sigma; h^t) := \frac{1-\delta}{1-\delta^{T-t}} \sum_{s=t}^{T-1} \delta^{s-t} \mathbb{E}_\sigma \left[u(A^s) \,\middle|\, H^t = h^t \right]$$

or, equivalently, $U(\sigma|_{h^t})$ in the $(T-t)$-repeated game.

Note:

- Infinitely repeated games are particularly tractable since every continuation game is identical to the entire game.

Decomposition of the Continuation Value

Decomposition of the continuation value:

- For any strategy profile σ and any history h^t,

$$
\begin{aligned}
U(\sigma; h^t) &= \frac{1-\delta}{1-\delta^{T-t}} \sum_{s=t}^{T-1} \delta^{(s-t)} \mathbb{E}_\sigma \left[u(A^s) \,\middle|\, H^t = h^t \right] \\
&= \frac{1-\delta}{1-\delta^{T-t}} \mathbb{E}_{\sigma(h^t)} \left[u(A^t) \right] + \frac{1-\delta}{1-\delta^{T-t}} \delta \sum_{s=t+1}^{T-1} \delta^{s-t-1} \mathbb{E}_\sigma \left[u(A^s) \,\middle|\, H^t = h^t \right] \\
&= \underbrace{\frac{1-\delta}{1-\delta^{T-t}}}_{=1-\delta_{T-t}} \mathbb{E}_{\sigma(h^t)} \left[u(A^t) \right] + \underbrace{\frac{\delta(1-\delta^{T-t-1})}{1-\delta^{T-t}}}_{=:\,\delta_{T-t}} \mathbb{E}_{\sigma(h^t)} [U(\sigma|_{h^t A^t})].
\end{aligned}
$$

- The relative weight δ_{T-t} of the continuation value is increasing from $\delta_0 = 0$ to $\delta_\infty = \delta$ in the length of the continuation game.
- Idea: we can use continuation strategies after histories $h^t a^t$ to support non-static Nash play in period t.

Intertemporal Incentives

Intertemporal incentives:

- Players are willing to play $\sigma(h^t)$ after history h^t if and only if

$$(1-\delta_{T-t})\mathbb{E}_{\sigma(h^t)}\left[u_i(A^t)\right] + \delta_T \mathbb{E}_{\sigma(h^t)}[U_i(\sigma|_{h^t A^t})]$$

$$\geq (1-\delta_{T-t})\mathbb{E}_{a_i,\sigma_{-i}(h^t)}\left[u_i(A^t)\right] + \delta_T \mathbb{E}_{a_i,\sigma_{-i}(h^t)}[U_i(\sigma|_{h^t A^t})].$$

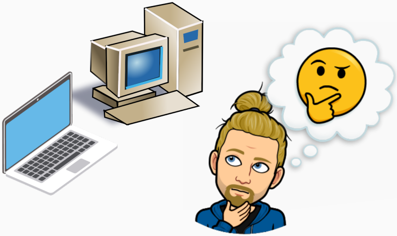
for every $a_i \in \mathcal{A}_i$ and every player i.

Credible punishments/rewards:

- In any SPE σ, we know that $\sigma|_{h^t a^t}$ is an SPE of the continuation game after $h^t a^t$, hence the rewards/punishments are credible.
- To construct an SPE that supports play of α, we need to find reward/punishment SPEs $\sigma|_{h^t a^t}$ that satisfy support play of α.

Product-Choice Game

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



Static game:

- The Firm can put (H)igh of (L)ow effort into its product.
- The Consumer chooses a (H)igh- or a (L)ow-priced product.
- The unique static Nash equilibrium is (L, L).

Subgame-perfect equilibria:

- Repeated play of (L, L) is an SPE by Lemma 11.1.
- Is there an SPE, in which the high-quality product is produced?

Uniqueness of SPE

Lemma 11.3

In a finitely repeated game with a unique static Nash equilbirium α_N, the unique SPE is repeated play of α_N.

Proof by induction:

- Fix any SPE σ and observe that $\sigma(h^T) = \alpha_N$ by subgame perfection.
- Inductive hypothesis: $\sigma|_{h^{t+1}}$ is repeated play of α_N for any $h^{t+1} \in \mathcal{H}^{t+1}$.
- For any history h^t, and any SPE $\widehat{\sigma}|_{h^t}$, we have

$$U(\widehat{\sigma}|_{h^t}) = (1-\delta_{T-t})\mathbb{E}_{\widehat{\sigma}|_{h^t}}\left[u_i(A^t)\right] + \delta_T \underbrace{\mathbb{E}_{\widehat{\sigma}|_{h^t}}\left[U_i(\sigma|_{h^t A^t})\right]}_{=u_i(\alpha_N)}.$$

- Therefore, $\widehat{\sigma}|_{h^t}$ plays α_N in every period.

Climate Agreement

	A	M	V
A	$4, 4$	$1.5, 4.5$	$-1, 5$
M	$4.5, 1.5$	$2, 2$	$0, 1$
V	$5, -1$	$1, 4$	$0, 0$

Modification:

- In the Paris Agreement, each country can set its own climate goals.
- Suppose each country can either set ambitious goals (A), moderate goals (M), or violate (V) the agreement.
- The stage game has two pure Nash equilibria: (V, V) and (M, M).
- Are there non-trivial SPE in a twice repeated game?
- Is it possible to support (A, A) in the first period?

Supporting Ambitious Goals

Punishments/rewards:

- Last period play has to be either (V, V) or (M, M).
- Idea: reward players with (M, M) for playing (A, A) and use (V, V) as a punishment if anything else has been played.

Deviations:

- The most profitable deviation is to play V in the first period.
- This deviation is profitable if and only if

$$\frac{1-\delta}{1-\delta^2}(5+0) > \frac{1-\delta}{1-\delta^2}(4+2\delta).$$

- Therefore, $\sigma(h^0) = (A, A)$, $\sigma(A, A) = (M, M)$, and $\sigma(h^1) = (V, V)$ for $h^1 \neq (A, A)$ is a subgame-perfect equilibrium if $\delta \geq \frac{1}{2}$.

Intertemporal Incentives

Key ingredients:

- At least two continuation equilibria with non-identical payoffs:
  - One “reward equilibrium” for following the equilibrium strategy,
  - One “punishment equilibrium” for deviating.
- Sufficiently high discount factor so that the difference in continuation payoffs outweighs potential profits in the current period.

Problem in longer games:

- We need to verify that no player has a profitable deviation.
- The number of possible deviations to consider grows exponentially with the length of the continuation game.

The One-Shot Deviation Principle

Theorem 11.4 (One-shot deviation principle)

A strategy profile σ in a repeated game is subgame perfect if and only if no player i has a profitable one-shot deviation, i.e., a deviation $\widetilde{\sigma}_i$ such that $\widetilde{\sigma}_i(h) = \sigma_i(h)$ for all but a single history $h \in \mathcal{H}$.

Proof:

- Necessity follows since an SPE has no profitable deviations, hence also no profitable one-shot deviations.
- For sufficiency, suppose that σ is not subgame perfect.
- There exists a history h^t, a player i, and a strategy $\widetilde{\sigma}_i$ of the continuation game after history h^t such that

$$U_i(\widetilde{\sigma}_i, \sigma_{-i}|_{h^t}) > U_i(\sigma_i|_{h^t}, \sigma_{-i}|_{h^t}).$$

Proof of Theorem 11.4

Finite deviation:

- Let $\varepsilon := U_i(\widetilde{\sigma}_i, \sigma_{-i}|_{h^t}) - U_i(\sigma|_{h^t}) > 0$ denote player i's gain from the deviation and let T be large enough that $\delta^T(M - m) < \frac{\varepsilon}{2}$, where

$$m = \min_{i,a} u_i(a), \qquad M = \max_{i,a} u_i(a).$$

- Deviation after time $t + T$ has a payoff impact of at most $\frac{\varepsilon}{2}$.
- Define a T-period deviation $\widehat{\sigma}_i$ by setting

$$\widehat{\sigma}_i(\widetilde{h}^s) = \begin{cases} \widetilde{\sigma}_i(\widetilde{h}^s) & \text{if } s \leq T, \\ \sigma_i|_{h^t}(\widetilde{h}^s) & \text{if } s > T. \end{cases}$$

- Triangle inequality implies that $\widehat{\sigma}_i$ is a profitable deviation with

$$U_i(\widehat{\sigma}_i, \sigma_{-i}|_{h^t}) - U_i(\sigma|_{h^t}) > \frac{\varepsilon}{2}.$$

Proof of Theorem 11.4

Does $\widehat{\sigma}_i$ already contain a one-shot deviation?

- Let $\widetilde{\mathcal{H}}^s$ denote the set of continuation histories $\widetilde{h}^s$ of h^t of length s that satisfy $\widehat{\sigma}_i(\widetilde{h}^s) \neq \sigma_i|_{h^t}(\widetilde{h}^s)$ and $P_{\widehat{\sigma}_i, \sigma_{-i}|_{h^t}}(\widetilde{h}^s) > 0$.
- Because $\widehat{\sigma}_i$ is a deviation in at most T-periods, there exists a time $\tau > 0$ such that $\widetilde{\mathcal{H}}^\tau \neq \emptyset$ and $\widetilde{\mathcal{H}}^s = \emptyset$ for all $s > \tau$.
- This means $\widehat{\sigma}_i$ differs from $\sigma_i|_{h^t}$ only in the first τ periods after h^t. Call $\widehat{\sigma}_i$ a τ-period deviation.
- Suppose first that there exists $\widetilde{h}^\tau \in \widetilde{\mathcal{H}}^\tau$ with

$$U_i\left(\widehat{\sigma}_i|_{\widetilde{h}^\tau}, \sigma_{-i}|_{h^t\widetilde{h}^\tau}\right) > U_i\left(\sigma|_{h^t\widetilde{h}^\tau}\right). \tag{2}$$

- Then $\widehat{\sigma}_i|_{\widetilde{h}^\tau}$ is a profitable one-shot deviation after $h^t\widetilde{h}^\tau$.

Proof of Theorem 11.4

If not, shorten the length of the deviation (part 1):

- If there exists no such history $\widetilde{h}^\tau \in \widetilde{\mathcal{H}}^\tau$, define the strategy $\bar{\sigma}_i$ by setting $\bar{\sigma}_i(\widetilde{h}^s) = \widehat{\sigma}_i(\widetilde{h}^s)$ if $s < \tau$ and $\sigma_i|_{h^t}(\widetilde{h}^s)$ otherwise.
- Since $\bar{\sigma}_i$ agrees with $\widehat{\sigma}_i$ for $\tau - 1$ periods, the two strategies induce the same distribution over histories in $\widetilde{\mathcal{H}}^\tau$, hence
  1. $P_{\bar{\sigma}_i,\sigma_{-i}|_{h^t}}\left(\widetilde{h}^\tau\right) = P_{\widehat{\sigma}_i,\sigma_{-i}|_{h^t}}\left(\widetilde{h}^\tau\right)$ for any $\widetilde{h}^\tau \in \widetilde{\mathcal{H}}^\tau$,
  2. $\mathbb{E}_{\bar{\sigma}_i,\sigma_{-i}|_{h^t}}\left[u_i(A^{t+s})\right] = \mathbb{E}_{\widehat{\sigma}_i,\sigma_{-i}|_{h^t}}\left[u_i(A^{t+s})\right]$ for any $s \leq \tau - 1$.
- Point 1. and the negation of (2) imply that

$$\mathbb{E}_{\bar{\sigma}_i,\sigma_{-i}|_{h^t}}\left[U_i(\bar{\sigma}_i|_{\widetilde{H}^\tau}, \sigma_{-i}|_{h^t\widetilde{H}^\tau})\right] = \mathbb{E}_{\bar{\sigma}_i,\sigma_{-i}|_{h^t}}\left[U_i(\sigma|_{h^t\widetilde{H}^\tau})\right] \tag{3}$$
$$\geq \mathbb{E}_{\widehat{\sigma}_i,\sigma_{-i}|_{h^t}}\left[U_i(\widehat{\sigma}_i|_{\widetilde{H}^\tau}, \sigma_{-i}|_{h^t\widetilde{H}^\tau})\right].$$

Proof of Theorem 11.4

If not, shorten the length of the deviation (part 2):

- Since the continuation value after history h^t is a convex combination of terms in point 2. and (3), it follows that

$$U_i(\bar{\sigma}_i, \sigma_{-i}|_{h^t}) \geq U_i(\widehat{\sigma}_i, \sigma_{-i}|_{h^t}).$$

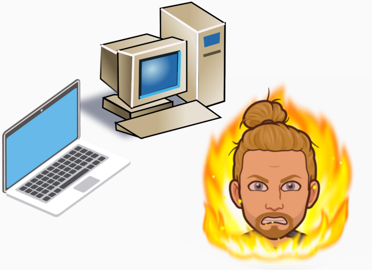
- Therefore, $\bar{\sigma}_i$ is a profitable $(\tau - 1)$-period deviation.

Proceed by backward induction:

- In the same way, either $\bar{\sigma}_i$ contains a profitable one-shot deviation or the last period of the deviation is not needed for $\bar{\sigma}_i$ to be profitable.
- By an iteration of this argument, we will eventually reach a profitable one-shot deviation.

Grim Trigger

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



Candidate strategy profile:

- The grim-trigger strategy profile σ is defined by

$$\sigma_i(h) = \begin{cases} E & \text{if } h = (H, H), \ldots, (H, H), \\ S & \text{otherwise.} \end{cases}$$

- Punishment continuation strategy profile is definitely subgame perfect.
- Is grim trigger an SPE in the infinitely repeated game?

Summary

Constructing equilibria:

- Time homogeneity and the one-shot deviation principle greatly simplify verifying that a given strategy profile is an SPE.
- We can construct SPEs by attaching punishments to unfavorable outcomes. This construction hinges on finding good punishment SPE.
- What do we do if the static Nash equilibrium is not a good punishment?

$\rightarrow$ We take Econ 8008 – Mircoeconomic Theory II.

Reputation games:

- Our knowledge of complete-information repeated games suffices.
- Next, we simplify the players' beliefs by introducing commitment types.

Check Your Understanding

True or false:

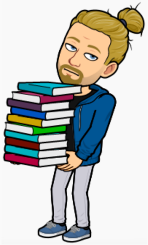
1. The board game Monopoly is a repeated game.
2. The one-shot deviation principle does not apply to extensive-form games with perfect information.
3. The one-shot deviation principle does not apply to Nash equilibria.

Scan me

4. In the 2-period modified climate agreement example, action profiles other than (A, A) can be played in an SPE during the first period.

Literature

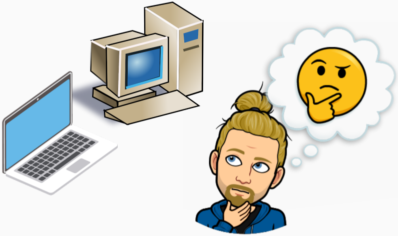
- D. Fudenberg and J. Tirole: Game Theory, Chapter 5, MIT Press, 1991
- G.J. Mailath and L. Samuelson: Repeated Games and Reputations: Long-Run Relationships, Chapter 2, Oxford University Press, 2006
- S. Tadelis: Game Theory: An Introduction, Chapter 10, Princeton University Press, 2013
- D. Abreu, 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: Efficiency and observability with long-run and short-run players, Journal of Economic Theory, **62** (1994), 103–135
- D. Fudenberg, D. K. Levine, and E. Maskin: The folk theorem with imperfect public information, Econometrica, **62** (1994), 997–1039



Reputation Games

Reputations?

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



Reputations?

- Grim trigger profile can be interpreted as maintaining a reputation for playing H: deviating destroys one's reputation forever.
- The set of continuation equilibria, however, is history-independent.
- Repeated play of H does not make the customer believe it is more likely that H will be played. There is no reputation effect.
- Consequently, no SPE supports H in a finitely repeated game.

Reputation Models

Reputation models:

- There is uncertainty about the quality/type of an institution.
- Individuals who interact with the institution report their experience and update social beliefs about the quality/type of the institution.
- Individuals are "short-lived", i.e., they interact with institution once.

Reputation effect:

- By repeatedly acting in a certain way, individuals must come to expect this behavior. The institution builds a reputation for acting that way.
- Institutions with a good reputation enjoy a benefit.
- Building a reputation may be costly in the short-run.

Newly Opened Restaurant

Updating of beliefs:

- Customers are unsure about the quality of the restaurant.
- Social beliefs are updated via Google/Yelp reviews.

Reputation effect:

- Building a good reputation is costly and it may require skilled personnel, developing unique recipes, good interior design, etc.
- If the restaurant manages to establish a good reputation, it receives more customers, or it can increase prices.

Electing Government Officials

Politician's type:

- Citizen's are unsure whether a politician is true to their word.
- Politician's past performance is an imperfect reflection about the politician's intentions and an indication of their future behavior.

Building a reputation:

- Politician works hard even when he/she is only appointed to low offices.
- If the politician succeeds in building a good reputation, he/she may be elected to a higher office that earns him/her many benefits.

Commitment Types

Definition 11.5

1. A commitment type is a type ϑ_c that invariably plays strategy $\sigma_1(\vartheta_c)$.
2. A commitment type ϑ_c is simple if there exists $\alpha_1 \in \Delta(\mathcal{A}_1)$ such that $\sigma_1(\vartheta_c, h^t) = \alpha_1$ for any history h^t. We also denote that type by ϑ_{α_1}.

Note:

- Commitment types do not maximize a utility function.
- In the product-choice game, a firm that is committed to delivering a high-quality product is a simple commitment type ϑ_H.
- Commitment types do not falter: they never deviate from their strategy profile, hence any deviations are attributed to payoff types.

Reputation Game

Player 1's types:

- We denote by Θ_c the set of player 1's possible commitment types.
- Additionally, player 1 may be a payoff type ϑ_p who maximizes

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

  where P_σ is defined from σ and prior beliefs $\mu_0 \in \Delta(\Theta)$ as usual.
- We denote by $\Theta = \Theta_c \cup \{\vartheta_p\}$ the set of all types.

Strategies:

- A strategy of Player 1 is a map $\sigma_1 : \Theta \times \mathcal{H} \to \Delta(\mathcal{A}_1)$.
- Type ϑ chooses a_1 after history h^t with probability $\sigma_1(\vartheta, h^t; a_1)$.
- It will be convenient to abbreviate $\sigma(\vartheta_p) = (\sigma_1(\vartheta_p), \sigma_2)$.

Updating of Beliefs

Beliefs:

- Let $\mu(h^t) \in \Delta(\Theta)$ denote player 2's beliefs about θ after history h^t, which assigns probability $\mu(h^t; \vartheta) = P_\sigma(\theta = \vartheta \,|\, H^t = h^t)$ to type ϑ.
- Player 2's belief process $(\mu_t)_{t \geq 0}$ is defined by $\mu_t = \mu(H^t)$.
- Observe that $(\mu_t)_{t \geq 0}$ is a sequence of random variables.

Updating of beliefs:

- After observing history $h^{t+1} = h^t a^t$ with $P_\sigma\left(h^{t+1}\right) > 0$, the beliefs are

$$\mu(h^{t+1}; \vartheta) = \frac{\sigma(\vartheta, h^t; a^t)\mu(h^t; \vartheta)}{\sum_{\vartheta' \in \Theta} \sigma(\vartheta', h^t; a^t)\mu(h^t; \vartheta')}.$$

- After any other history, the beliefs are update to $\mu(h^{t+1}) = \delta_{\vartheta_p}$.
- Consequence: any strategy profile leads to well-defined off-path beliefs.

Short-Lived Players

Player 2 is short-lived:

- Each instance of player 2 "lives" for only one round.
- Therefore, player 2 plays a stage-game best reply to $\mu(h^t)$.
- For any action α_1, denote player 2's stage-game best responses by

$$\mathcal{B}_2(\alpha_1) := \left\{\alpha_2 \in \Delta(\mathcal{A}_2) \,\middle|\, u_2(\alpha_1, \alpha_2) = \max_{a_2 \in \mathcal{A}_2} u_2(\alpha_1, a_2)\right\}.$$

- Given beliefs $\mu(h^t)$, player 2 expects to see $\alpha_1^{\sigma,\mu}(h^t)$ defined by

$$\alpha_1^{\sigma,\mu}(h^t; a_1) = \sum_{\vartheta \in \Theta} \sigma_1(\vartheta, h^t; a_1)\mu(h^t; \vartheta).$$

- Player 2 plays an action α_2 from $\mathcal{B}_2(\alpha_1^{\sigma,\mu}(h^t))$, which maximizes

$$\mathbb{E}_\sigma\left[u_2(\sigma_1(\theta, h^t), \alpha_2) \,\middle|\, H^t = h^t\right] = \sum_{\vartheta \in \Theta} u_2(\sigma_1(\vartheta, h^t), \alpha_2)\mu(h^t; \vartheta).$$

Game Dynamics and Equilibrium

Definition 11.6

Strategy profile σ is a Nash equilibrium of the reputation game if for all $\widehat{\sigma}_1$, $U_1(\sigma) \geq U_1(\widehat{\sigma}_1, \sigma_2)$ and for all h^t with $P_\sigma(H^t = h^t) > 0$ and all $a_2 \in \mathcal{A}_2$,

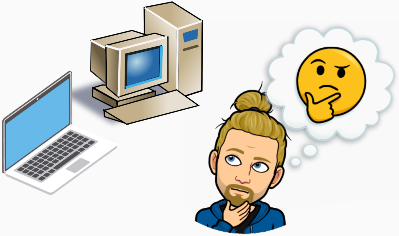
$$\mathbb{E}_\sigma\left[u_2(\sigma_1(h^t, \theta), \sigma_2(h^t)) \,\middle|\, H^t = h^t\right] \geq \mathbb{E}_\sigma\left[u_2(\sigma_1(h^t, \theta), a_2) \,\middle|\, H^t = h^t\right].$$

Leader-follower dynamic:

- Player 2 is short-lived and maximizes myopically, given his/her beliefs.
- Player 1 anticipates player 2's best response and optimizes accordingly.
- Future threats or rewards have no impact on player 2's behavior, hence player 1 essentially faces a single-player optimization problem.

Two-Stage Product-Choice Game

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



Firm has two types:

- Commitment type ϑ_H is committed to playing H in both periods.
- Payoff type maximizes the sum of discounted payoffs.

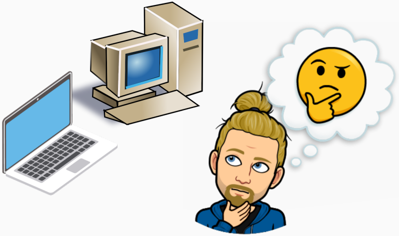
Equilibrium in second period:

- Payoff type must choose L in the second period because it is dominant.
- In period 2, Customers best responds with H if and only if

$$3\mu(h^1;\vartheta_H) \geq 2\mu(h^1;\vartheta_H) + 1 - \mu(h^1;\vartheta_H) \quad \Leftrightarrow \quad \mu(h^1;\vartheta_H) \geq \tfrac{1}{2}.$$

Two-Stage Product-Choice Game

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



Optimistic customers:

- If $\mu_0(\vartheta_H) \geq \frac{1}{2}$, Customers choose H in the first period.
- Mimicking the commitment type leads to $\mu(H; \vartheta_H) = \mu_0(\vartheta_H) \geq \frac{1}{2}$.
- Choosing H followed by L nets Firm $2 + 3\delta$.
- Choosing L twice yields $3 + \delta$, which is preferable only if $\delta \leq \frac{1}{2}$.

Pessimistic customers:

- If $\mu_0(\vartheta_H) < \frac{1}{2}$, Customers choose L in the first period.
- Is repeated play of (L, L) an equilibrium?

Two-Stage Product-Choice Game

Profitable deviation?

- Candidate profile $\sigma(h) = (L, L)$ for every history h.
- Consider a deviation by Firm to H in the first period.
- After observing H, Customers' beliefs in the second period are

$$\mu(H; \vartheta_H) = \frac{\mu_0(\vartheta_H)}{\mu_0(\vartheta_H) + \sigma_1(\vartheta_p, H)\mu_0(\vartheta_p)} = 1.$$

- Customer's best response in period 2 is to choose H.

Reputation effect:

- Type ϑ_p has an incentive to mimic type ϑ_H and play H in period 1.
- Building the reputation is costly: doing so nets Firm 0 in period 1.
- Having built the reputation, Firm can exploit it and earn 3 in period 2.
- If $3\delta \geq 1 + \delta$, i.e., $\delta \geq \frac{1}{2}$, then doing so is profitable.

Impact of Reputations

Power of reputations:

- In a finitely repeated game, repeated play of (L, L) is the only SPE.
- If there is a small possibility $\mu_0 > 0$ that the firm is committed to producing a high-quality product, (L, L) is no longer an equilibrium.
- Playing L is suboptimal because it fails to exploit the reputation effect.

Strong variation of beliefs:

- A typical ingredient in reputation arguments is the strong variation of beliefs after a deviation by the payoff type.
- Choosing H with certainty may not be part of an equilibrium. Nevertheless, any equilibrium must be robust to this ϑ_H-imitating deviation.

Question: Can we solve for an equilibrium?

Equilibrium in Behavior Strategies

Parametrizing Firm's Strategy:

- In any Nash equilibrium σ, we must have $\sigma_1(\vartheta_H, h) = H$ for any $h \in \mathcal{H}^0 \cup \mathcal{H}^1 = \{\emptyset, H, L\}$ and $\sigma_1(\vartheta_p, h^1) = L$ for any $h^1 \in \mathcal{H}^1$.
- Firm's strategy is entirely parametrized by $\sigma_1(\vartheta_p, \emptyset; H) = x$.

Parametrizing Customer's Strategy:

- Suppose Customers are pessimistic, i.e., $\mu_0(\vartheta_H) < \frac{1}{2}$.
- In any Nash equilibrium σ, we have $\sigma_2(\emptyset) \in \mathcal{B}_2(\mu_0 + x(1-\mu_0))$ and

$$\sigma_2(h^1) = \begin{cases} H & \text{if } \mu(h^1; \vartheta_H) \geq \frac{1}{2}, \\ L & \text{if } \mu(h^1; \vartheta_H) < \frac{1}{2}. \end{cases}$$

Parametrizing strategy profile:

- Strategy profile σ is parametrized by $x = \sigma_1(\vartheta_p, \emptyset; H) \Rightarrow$ write $P_\sigma = P_x$.

Equilibrium in Behavior Strategies

Parametrizing beliefs:

- $\mu(h) \in \Delta(\Theta)$ is a vector $(\mu(h;\vartheta_H), \mu(h;\vartheta_p))$.
- Identify $\mu(h) \cong \mu(h;\vartheta_H)$ and write $\mu(h;\vartheta_p) = 1 - \mu$.

Posterior as random variable:

- $\mu_1 := \mu(H^1)$ is a random variable, taking values

$$\mu(H) = \frac{\sigma(\vartheta_H, \emptyset; H)\mu_0}{\sigma(\vartheta_H, \emptyset; H)\mu_0 + \sigma(\vartheta_p, \emptyset; H)(1-\mu_0)} = \frac{\mu_0}{\mu_0 + x(1-\mu_0)},$$

$$\mu(L) = \frac{\sigma(\vartheta_H, \emptyset; L)\mu_0}{\sigma(\vartheta_H, \emptyset; L)\mu_0 + \sigma(\vartheta_p, \emptyset; L)(1-\mu_0)} = 0,$$

  on the events $\{H^1 = H\}$ and $\{H^1 = L\}$, respectively.
- It follows that $P_x(\mu_1 = \mu(H)) = x$ and $P_x(\mu_1 = 0) = 1 - x$.

Equilibrium in Behavior Strategies

Continuation value:

- The strategic type maximizes

$$\begin{aligned} U_1(x) &= \mathbb{E}_x\left[u_1(A_0^1, A_0^2)\right] + \delta\mathbb{E}_x\left[u_1(L, A_1^2)\right] \\ &= (1-x) + 2\cdot 1_{\{\mu_0 + x(1-\mu_0)\geq 0.5\}} + \delta\big(1 + 2P_x(\mu_1 \geq 0.5)\big) \\ &= 1 + \delta + 2\cdot 1_{\{\mu_0 + x(1-\mu_0)\geq 0.5\}} - x + 2\delta x 1_{\{\mu(H)\geq 0.5\}}. \end{aligned}$$

- Observe that $\mu(H) \geq \frac{1}{2}$ if and only if $x \leq \frac{\mu_0}{1-\mu_0}$.
- Moreover, $\mu_0 + x(1-\mu_0) \geq \frac{1}{2}$ if and only if $x \geq \frac{1-2\mu_0}{2(1-\mu_0)}$.
- $U_1(x)$ has a discontinuities at $x_0 := \frac{\mu_0}{1-\mu_0}$ and $x_1 := \frac{1-2\mu_0}{2(1-\mu_0)}$.
- Observe that $x_0 \in [0,1)$ and $x_1 \in \left(0, \frac{1}{2}\right]$. Moreover, $x_0(\mu_0)$ is increasing, $x_1(\mu_0)$ is decreasing with $x_1 \leq x_0$ if and only if $\mu_0 \geq \frac{1}{4}$.

Equilibrium in Behavior Strategies

Solving for maximum:

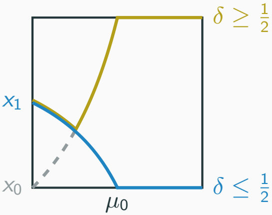
- Due to the discontinuities, U_1 is maximized either at $x \in \{0, x_0, x_1, 1\}$ or where $\frac{\partial U_1(x)}{\partial x} = -1 + 2\delta 1_{\{x \le x_0\}} = 0$.
- Derivative is 0 only if $\delta = \frac{1}{2}$ and $x \le x_0$. In that case, if x_0 maximizes Firm's average discounted payoff, then so does $x \in [x_1 1_{\{x_1 \le x_0\}}, x_0]$.
- We compare $U_1(0) = 1 + \delta$, $U_1(1) = 2 + \delta$, and

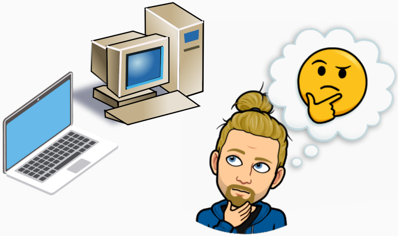
$$U_1(x_0) = 1 + \delta + (2\delta - 1)x_0 + 2 \cdot 1_{\{x_1 \le x_0\}},$$

$$U_1(x_1) = 3 + \delta - x_1 + 2\delta x_1 1_{\{x_1 \le x_0\}}.$$

- Clearly, $U_1(x_1) > U_1(1) > U_1(0)$.
- If $\mu_0 \ge \frac{1}{4}$, then $U_1(x_1) \ge U_1(x_0)$ if and only if $\delta \le \frac{1}{2}$.
- If $\mu_0 < \frac{1}{4}$, then $U_1(x_1) \ge U_1(x_0)$.

Equilibrium in Behavior Strategies





Solving for maximum:

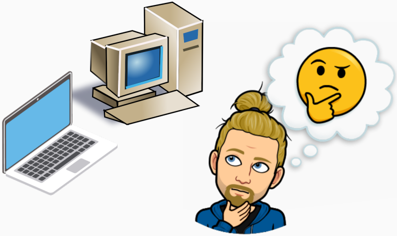
- We conclude that U_1 is maximized at

$$x_* \in \begin{cases} \{x_1\} & \text{if } \mu_0 < \frac{1}{4} \text{ or } \mu_0 \geq \frac{1}{4} \text{ and } \delta < \frac{1}{2}, \\ [x_1, x_0] & \frac{1}{4} \geq \mu_0 \text{ and } \delta = \frac{1}{2}, \\ \{x_0\} & \frac{1}{4} \geq \mu_0 \text{ and } \delta > \frac{1}{2}. \end{cases}$$

- x_0 and x_1 are the cutoffs, above which Customers choose L and H in the first period, respectively, and H in the second period if $A_1^1 = H$.

Infinitely Repeated Product-Choice Game

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



Infinitely repeated game:

- Is repeated play of (L, L) is an equilibrium?
- Consider deviation $\widehat{\sigma}_1$ that plays H after every history.
- Again, Customers are convinced they are facing ϑ_H after first period.
- Deviation yields 0 in the first period, but a continuation payoff of 2.
- Deviation is profitable if $\delta \geq \frac{1}{2}$.
- Patient players strictly benefit from exploiting the reputation effect.

Check Your Understanding

True or false:

1. There is no reputation effect in standard repeated games because the continuation game is equivalent to the original game.
2. If the long-lived player has multiple payoff types, then it is insufficient to consider Nash equilibria.

Scan me

3. Short-lived players have a unique best response for any beliefs.
4. In a 2-action, 2-period reputation game, player 1's best response in the first period is attained either at the boundary, at a discontinuity of U_1, or where $U_1' = 0$.

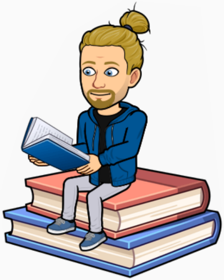
Literature

- G.J. Mailath and L. Samuelson: Repeated Games and Reputations: Long-Run Relationships, Chapter 15, 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



Reputation Effect with Perfect Monitoring

Reputation Effect

Evolution of beliefs:

- We know how player 2 updates beliefs about player 1's type.
- However, player 2 does not care whether he/she faces a commitment type or a strategic type that mimics the commitment type.
- More relevant: what are player 2's beliefs about expected play?

Outcomes of the perfect-monitoring game:

- Let Ω be the set of all possible outcomes $\omega = \left(\vartheta, a^0, a^1, \ldots\right)$.
- Outcome of the game (θ, A) is an Ω-valued random variable.
- Want to exploit the correlation of A and θ to express player 2's beliefs about expected continuation play.
- Player 2 cannot observe θ: A_2^t is h^t-conditionally independent of θ.

Expectation of Future Play

Reputation effect:

- In the product-choice game, it was beneficial for the strategic player to mimic the commitment type ϑ_H.
- Fix an action $\widehat{a}_1 \in \mathcal{A}_1$ and a simple commitment type $\vartheta_{\widehat{a}_1} \in \Theta_c$.
- Does player 2 come to expect $\widehat{a}_1$ after repeated play of $\widehat{a}_1$?

Expectation of future play:

- In period t, player 2 expects to see $\widehat{a}_1$ with a likelihood

$$q_t(h^t) := P_\sigma\left(A_1^t = \widehat{a}_1 \,\middle|\, H^t = h^t\right) = \sum_{\vartheta \in \Theta} \sigma_1(\vartheta, h^t; \widehat{a}_1)\mu(h^t; \vartheta).$$

- We refer to $q_t(H^t)$ simply by q_t. Note that q_t is a random variable: before time t, it is unknown which h^t will materialize.

Mimicking a Commitment Type

Mimicking a commitment type:

- Let $\Omega' := \{\omega \in \Omega \mid A_1^t(\omega) = \widehat{a}_1 \text{ for all } t\}$ denote the set of outcomes, in which $\widehat{a}_1$ is played in every period.
- Ω' contains outcomes $\omega = (\vartheta, a^0, a^1, \ldots)$ with:
  - Player 1 is the commitment type $\vartheta_{\widehat{a}_1}$,
  - Player 1 is the strategic type that imitates $\vartheta_{\widehat{a}_1}$,
  - Player 1 is any other commitment type (or a strategic type) that plays $\widehat{a}_1$ with positive probability in every period and $\widehat{a}_1$ was always realized.

How does expectation about future play evolve on Ω'?

- Intuition: q_t is increasing on Ω'.
- However, this intuition need not be correct.

Evolution of q_t in the Product-Choice Game

Commitment types:

- Simple commitment type ϑ_H.
- Commitment type ϑ_t plays H in period $s < t$ and L after.
- Let $\widehat{a}_1 = H$, that is, $\Omega' = \{\omega \in \Omega \mid A_1^t(\omega) = H\}$.

Evolution of q_t:

- Consider the strategy profile $\sigma(\vartheta_p, h^t) = (H, h)$ for any h^t.
- Then $q_0 = 1 - \mu_0(\vartheta_0)$ and

$$q_1(H, h) = \frac{1 - \mu_0(\vartheta_0) - \mu_0(\vartheta_1)}{1 - \mu_0(\vartheta_0)}.$$

- Note that $q_1(H, h) < q_0$ if $\mu_0(\vartheta_1) > \mu_0(\vartheta_0)(1 - \mu_0(\vartheta_0))$.

Whether q_t increases or decreases depends not only on σ, but also on prior beliefs and on commitment types that are being ruled out.

Reputation Effect

Lemma 11.7

For any $\zeta \in [0,1)$, let $n_\zeta := |\{t \mid q_t \leq \zeta\}|$. Fix $\widehat{a}_1 \in \mathcal{A}_1$ with $\vartheta_{\widehat{a}_1} \in \Theta_c$ and suppose that $\mu_0(\vartheta_{\widehat{a}_1}) \in [\mu_*, 1)$ for some $\mu_* > 0$. For any profile σ,

$$P_\sigma\left(n_\zeta > \frac{\ln \mu_*}{\ln \zeta} \,\middle|\, \Omega'\right) = 0.$$

Moreover, $\mu_t(\vartheta_{\widehat{a}_1})$ is non-decreasing for P_σ-almost every $\omega \in \Omega'$.

- If payoff type mimics commitment type $\vartheta_{\widehat{a}_1}$, player 2 must expect $\widehat{a}_1$ with large probability in all but finitely many periods.
- Bound $\frac{\ln \mu^*}{\ln \zeta}$ on number of those periods does not depend on σ.
- $\mu_t(\vartheta_{\widehat{a}_1})$ may not converge to 1: if $\sigma(\vartheta_p)$ mimics the commitment type, player 2 places positive beliefs on facing a $\vartheta_{\widehat{a}_1}$-imitating payoff type.

Proof: Step 1

Monotonicity of μ:

- Fix σ and $h^{t+1} = (h^t, a^t)$ with $a_1^t = \widehat{a}_1$ and $P_\sigma\big(H^{t+1} = h^{t+1}\big) > 0$.
- Recall $q(h^t) = P_\sigma(A_1^t = \widehat{a}_1 \,|\, H^t = h^t)$. Bayes' rule implies that

$$P_\sigma\big(\theta = \vartheta_{\widehat{a}_1} \,\big|\, h^{t+1}\big) = \frac{P_\sigma(A^t = a^t \,|\, \vartheta_{\widehat{a}_1}, h^t) P_\sigma(\theta = \vartheta_{\widehat{a}_1} \,|\, h^t)}{P_\sigma(A^t = a^t \,|\, h^t)}$$

$$= \frac{P_\sigma(A_2^t = a_2^t \,|\, \vartheta_{\widehat{a}_1}, h^t)\mu(h^t; \vartheta_{\widehat{a}_1})}{q_t(h^t)\sigma_2(h^t; a_t^2)} = \frac{\mu(h^t; \vartheta_{\widehat{a}_1})}{q(h^t)}.$$

- Since $q(h^t) \leq 1$, this implies that $\mu(h^{t+1}; \vartheta_{\widehat{a}_1}) \geq \mu(h^t; \vartheta_{\widehat{a}_1})$.
- Moreover, for every $\omega = (\vartheta, a^0, a^1, \ldots)$ in Ω' with $P_\sigma((\theta, A) = \omega) > 0$, history $(a^0, \ldots, a^t)$ of any length $t + 1$ is of the above form.
- This shows that $\mu_t(\vartheta_{\widehat{a}_1})$ is non-decreasing for P_σ-a.e. $\omega \in \Omega'$.

Proof: Step 2

Finding a bound for ζ:

- Fix $\omega = (\vartheta, a^0, a^1, \dots) \in \Omega'$ with $P_\sigma((\theta, A) = \omega) > 0$ and denote by $h^t = (a^0, \dots, a^{t-1})$ the associated histories.
- From previous slide: $\mu(h^t; \vartheta_{\widehat{a}_1}) = q(h^t)\mu(h^{t+1}; \vartheta_{\widehat{a}_1})$ for all t.
- Since $\mu_0(\vartheta_{\widehat{a}_1}) \geq \mu^* > 0$ by assumption, we obtain

$$\mu^* \leq \mu_0(\vartheta_{\widehat{a}_1}) = q(h^0)\mu(h^1; \vartheta_{\widehat{a}_1}) = q(h^0)q(h^1)\mu(h_2; \vartheta_{\widehat{a}_1})$$

$$= \dots = \left(\prod_{s=0}^{t-1} q(h^s)\right)\mu(h^t; \vartheta_{\widehat{a}_1}) \leq \prod_{s=0}^{t-1} q(h^s).$$

- Taking the limit as $t \to \infty$ yields $\mu^* \leq \prod_{s=0}^{\infty} q(h^s)$, hence

$$\mu_* \leq \prod_{s=0}^{\infty} q(h^s) \leq \prod_{\{s: q_s \leq \zeta\}}^{\infty} q(h^s) \prod_{\{s: q_s > \zeta\}}^{\infty} q(h^s) \leq \prod_{\{s: q_s \leq \zeta\}}^{\infty} q(h^s) \leq \zeta^{n_\zeta(\omega)}.$$

Proof: Conclusion

Conclusion:

- We have deduced $\mu_* \leq \zeta^{n_\zeta(\omega)}$ for P_σ-almost every $\omega \in \Omega'$, hence

$$P_\sigma\big(\mu_* \leq \zeta^{n_\zeta} \,\big|\, \Omega'\big) = 1.$$

- Therefore,

$$P_\sigma\left(n_\zeta > \frac{\ln \mu^*}{\ln \zeta} \,\middle|\, \Omega'\right) = P_\sigma\big(\ln\big(\zeta^{n_\zeta}\big) < \ln \mu^* \,\big|\, \Omega'\big) = 0.$$

Exploiting the Reputation Effect

Reputation effect:

- Imitating a simple commitment type $\vartheta_{\widehat{a}_1}$ leads player 2 to expect $\widehat{a}_1$ with high probability ζ in all but finitely many periods.
- Choosing ζ sufficiently high will cause player 2 to best reply to $\widehat{a}_1$.

Which commitment type should player 1 imitate?

- If player 1 could commit to playing $a_1 \in \mathcal{A}_1$, he/she would get:

$$\underline{v}_1(a^1) = \min_{a^2 \in B(a^1)} u_1(a^1, a^2).$$

- Player 1 best imitates commitment type $\vartheta_{\widehat{a}_1}$ that maximizes $\underline{v}_1(\widehat{a}_1)$.

Reputation Bound

Theorem 11.8

Let $\widehat{\mathcal{A}}_1 \subseteq \mathcal{A}_1$ denote the set of player 1's pure actions a_1, for which there is a simple commitment type ϑ_{a_1} with $\mu_0(\vartheta_{a_1}) > 0$. Suppose $\mathcal{A}_2$ is finite and $\mu_0(\vartheta_p) > 0$. Then there exists $K(\mu_0)$ such that

$$\inf_{Eq.\ \sigma} U_1(\sigma) \geq \delta^K \max_{a_1 \in \widehat{\mathcal{A}}_1} \underline{v}_1(a_1) + (1 - \delta^K) \min_{a \in \mathcal{A}} u_1(a).$$

Interpretation:

- Player 1 builds a reputation to be his preferred commitment type $\widehat{a}_1$.
- Reputation effect: in all but $K(\mu_0)$ periods, player 2 best replies to $\widehat{a}_1$.
- Trade-off: building a reputation is costly, but yields player 1 his/her highest-possible commitment payoff in the long-run.

Note: Bound applies to *any* Nash equilibrium payoff.

Proof: Step 1

Choose preferred commitment type:

- Let $\widehat{a}_1$ be player 1's preferred simple commitment type:

$$\widehat{a}_1 \in \arg\max_{a_1 \in \widehat{\mathcal{A}}_1} \underline{v}_1(a_1).$$

- Let $\mu^* := \mu_0(\widehat{a}_1)$. Note that by assumption $\mu^* > 0$.

Choose ζ such that player 2 best replies to $\widehat{a}_1$:

- Because $\mathcal{A}_2$ is finite, there exists $\zeta \in (0,1)$ such that if $\alpha_1(\widehat{a}_1) > \zeta$.

$$\mathcal{B}_2(\alpha_1) \subset \mathcal{B}_2(\widehat{a}_1).$$

- Fix a Nash equilibrium σ. Then for all h^t, for which $q(h^t) > \zeta$,

$$\mathcal{B}_2\left(\mathbb{E}_\sigma\left[\sigma_1^t(h^t,\theta) \,\middle|\, h^t\right]\right) \subset \mathcal{B}_2(\widehat{a}_1).$$

Proof: Step 2

Apply reputation effect:

- Set $K = \frac{\ln \mu(\vartheta_{\widehat{a}_1})}{\ln \zeta}$.
- Lemma 11.7 implies that, conditional on player 1 always playing $\widehat{a}_1$, $q_t \leq \zeta$ for no more than K periods t with P_σ-probability 1.
- Let $\widehat{\sigma}_1$ be the deviation that always plays $\widehat{a}_1$. It follows that

$$P_{\widehat{\sigma}_1,\sigma_2}(\{t \mid q_t \leq \zeta\} < K) = P_\sigma(\{t \mid q_t \leq \zeta\} < K \,|\, \vartheta_{\widehat{a}_1}) = 1.$$

- Since σ_1 has no profitable deviations

$$U_1(\sigma) \geq U_1(\widehat{\sigma}_1, \sigma_2) \geq \delta^K \max_{a^1 \in \widehat{\mathcal{A}}_1} \underline{v}_1(a^1) + (1-\delta^K) \min_{a \in \mathcal{A}} u_1(a).$$

Patient Long-Lived Player

Corollary 11.9

Let $\vartheta_{\hat{a}_1}$ be player 1's preferred simple commitment type with $\mu_0(\vartheta_{\hat{a}_1}) > 0$. Suppose that $\mathcal{A}_2$ is finite and $\mu_0(\vartheta_p) > 0$. For any $\varepsilon > 0$, there exists $\underline{\delta} \in (0,1)$ such that for any $\delta > \underline{\delta}$,

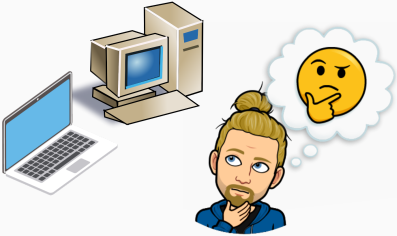
$$\inf_{Eq.\ \sigma} U_1(\sigma) \geq \max_{a_1 \in \mathcal{A}_1} \underline{v}_1(a_1) - \varepsilon.$$

As long-lived player gets patient:

- Building a reputation becomes cheaper.
- As $\delta \to 1$ reputation effect dominates all other considerations and only a single equilibrium payoff remains.

Reputation Bound in Product-Choice Game

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



Reputation bound:

- As long as there exists a commitment type ϑ_H with $\mu_0(\vartheta_H) > 0$, Theorem 11.8 establishes a lower bound for all Nash equilibria.
- Lower bound on equilibrium is increasing in δ because "building the reputation" is comparatively cheaper.
- By Corollary 11.9, payoff in any equilibrium converges to 2 as $\delta \to 1$.

Summary

Reputation effect:

- If the strategic type plays a commitment action $\widehat{a}_1$ repeatedly, the short-lived player must come to expect it eventually.
- Specifically, in all but finitely many periods, player 2 will best respond to the commitment action $\widehat{a}_1$.

Patient players:

- In *every* equilibrium, a reputation is built for the best commitment type.
- Building a reputation is costly in the short-run, but not doing so is suboptimal because it fails to exploit the reputation effect.
- Reputation effect is much stronger in an infinitely repeated game than in the twice repeated game.

Literature

- G.J. Mailath and L. Samuelson: Repeated Games and Reputations: Long-Run Relationships, Chapter 15, Oxford University Press, 2006
- S. Tadelis: Game Theory: An Introduction, Chapter 17, Princeton University Press, 2013
- D. Fudenberg, and D. K. Levine: Reputation and Equilibrium Selection in Games with a Patient Player, Econometrica, **57** (1989), 759–778