Maintaining a Reputation when Strategies are Imperfectly Observed

Fudenberg and Levine (ReStud, 1992) summary by N. Antić

In a repeated game players can develop a reputation for playing in a specific way. Building a reputation can take time, so patient players are more likely to invest.

Example

- The main point of this paper can be illustrated in a repeated "Chain Store Paradox" example

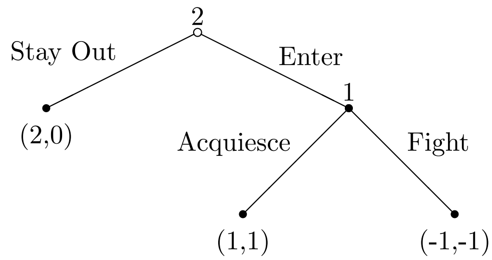

Figure 1: Stage game of the chain store paradox

- Monopolist facing an infinite sequence of potential entrants, can respond aggressively or passively
  - Period t entrant observes the entire preceding history
- Assume the monopolist can be a commitment type with a preference for fighting
  - All entrants have a common prior about this, $\varepsilon > 0$

Theorem. *In any sequential equilibrium, if δ is close to 1, then player 1's expected average payoff in equilibrium is close to 2.*

Proof Sketch. Fix a sequential equilibrium and let t be the first period that player 1 plays "Acquiesce". If $t = \infty$, player 2 is playing "Stay Out" and player 1 gets payoff 2. If $t < \infty$, then deviating to playing "Fight" in all periods will give payoff at least -1 in period t and say s periods after it (until player 1's posterior is sufficiently high) and payoff 2 in subsequent periods. □

- Result extends to finite number of commitment types

The Basic Environment

- 2 Players in each period, player 1 (the long-run player) and Player 2 (one of a sequence of short-run players)
  - Denote short-run player in period t by player 2_t
- Stage game pure action sets, A_1 and A_2, are finite (not critical)
  - Denote mixed actions by $\alpha_i \in \Delta(A_i)$
- Imperfect public monitoring
  - Players observe a random outcome $y \in Y$, where $|Y| = M < \infty$
  - Given action profile $a \in A$, the probability of signal y is $\rho(y|a)$
- Includes perfect monitoring as a special case
  - Another special case is an extensive form stage game where only terminal payoffs are observable
- All short-run players have a single type
- Payoff of player 2 is common knowledge
  - Depends only on public signal y and not directly on a_1
  - Same assumption for long-run players
- Short-run players' vNM utility index is $u_2 \colon Y \times A_2 \to \mathbb{R}$
  - Player 2's expected payoff from mixed action $\alpha \in \Delta(A)$ is

$$v_2(\alpha) = \sum_{(a_1,a_2)\in A} u_2(y, a_2)\, \rho(y|\,(a_1, a_2))\, \alpha_1(a_1)\, \alpha_2(a_2)$$

- Player 1's type space, Ω, is a metric space
- Common knowledge that short-run players have identical prior, μ, about player 1's type
  - μ is a measure on $\mathcal{B}(\Omega)$
- Rational type, $\omega_0 \in \Omega$, has stationary preferences over time, with vNM utility index $u_1(\alpha_1, y, \omega_0)$
- Assume $\mu(\omega_0) > 0$
- Commitment types have a preference for playing a certain action—including mixed actions
  - Will need commitment types with preferences for all mixed actions
  - Trick to make them expected utility maximizers (non-stationary preferences over time)
- Utility index for player 1 is uniformly bounded: $u_1(\alpha_1, y, \omega) \in [\underline{u}, \overline{u}]$ for all ω
- Assume commitment types have full support
  - Let η be the measure on mixed actions induced by μ
  - By LDT write $\eta = \eta_0 + \eta_1$, where $\eta_0 \ll \lambda$
  - Assume Radon-Nykodym derivative of η_0 is bounded away from 0

Equilibrium

- History for player 2 is the public history $H_t \in Y^t$
- Pure strategy for player 2_t is $s_2^t \colon H_{t-1} \to A_2$
  - S_2^t denotes the set of all pure strategies for player 2_t
- Player 1 knows the public history and his private history $H_t^1 \in (A_1)^t$
- Pure strategy for player 1 in period t is $s_1 = \{s_1^t\}_{t=1}^{\infty}$ where $s_1^t \colon H_{t-1} \times H_{t-1}^1 \to A_1$
  - S_1 denotes the set of all pure strategies for player 1
- Mixed strategies for players 1 and 2_t are $\sigma_1 \in \Delta(S_1)$ and $\sigma_2^t \in \Delta(S_2^t)$, respectively
- A mixed strategy for player 2 is $\sigma_2 \in \Delta(\times_{t=1}^{\infty} S_2^t)$
- Mixed strategy profile $\sigma = (\sigma_1, \sigma_2) \in \Delta(S)$ induces a probability distribution over $\{a_1(t), a_2(t)\}_{t=1}^{\infty}$ and $\{y(t)\}_{t=1}^{\infty}$
- Let E_σ denote the expectation w.r.t. this distribution

- The average expected utility of player 1 is:

$$U_1(\sigma,\omega) = E_\sigma\left[(1-\delta)\sum_{t=1}^{\infty}\delta^{t-1}u_1(a_1(t), y(t), \omega)\right]$$

- Another way to think about mixed strategies, useful when type spaces are infinite
  - Developed by Milgrom and Weber (1985)
  - Requires Ω to be a Polish space
- A **distributional strategy** for player 1, $\mathscr{s}_1$, is a probability measure on Borel sets of $\Omega \times S_1$
  - Consistency requirement—the marginal distribution on Ω is μ
- Let $\mathscr{S}_1$ denote all the distributional strategies of player 1
- Note that for any $\Omega^+ \subset \Omega$, $\mathscr{s}_1(\Omega^+) \in \Delta(S_1)$ where

$$\mathscr{s}_1(\Omega^+)(s_1) = \mu(\Omega^+)^{-1}\int_{\Omega^+}\mathscr{s}_1(\omega, s_1)d\omega$$

- Short-run players can have incorrect beliefs about long-run player's strategy if outcomes are insufficient to identify actions
  - Related to self-confirming equilibrium in learning in games
- An action α_2 is an **ε-confirmed best response** to α_1 if (i) α_2 is not weakly dominated and (ii) there exists some α_1' such that:
  - $\alpha_2 \in \arg\max_{\alpha_2'} v_2(\alpha_1', \alpha_2')$
  - $\|\rho(\cdot|(\alpha_1,\alpha_2)) - \rho(\cdot|(\alpha_1',\alpha_2))\|_\infty < \varepsilon$
- Denote by $B_\varepsilon(\alpha_1)$ the set of ε-confirmed best responses to α_1
- $B_0(\alpha_1)$ is **not** the set of all undominated best responses
  - These are generalized best responses (Fudenberg and Levine, 1989)
- A **Nash equilibrium** is $(\mathscr{s}_1, \sigma_2) \in \mathscr{S}_1 \times \Delta(S_2)$ so that σ_2^t is a best response to $\mathscr{s}_1(\Omega)$ and $(\omega, s_1) \in supp(\mathscr{s}_1)$ implies s_1 is a best response to σ_2 by type ω
- Nash equilibrium exists
  - Existence in finite truncations of the game proven by Milgrom and Weber (1985)
  - Fudenberg and Levine (1983) show that for finite-action imperfect information games which are uniformly continuous mixed-strategy sequential equilibria exist
  - Action spaces and signal spaces are finite, U_1 and v_2 are uniformly continuous
- Let $\underline{N}_1(\delta,\omega)$ and $\overline{N}_1(\delta,\omega)$ be the inf and sup of type ω's payoff in any Nash equilibrium of the repeated game with discount rate δ
- Let ε-least commitment payoff for type ω be:

$$\underline{v}_1(\omega,\varepsilon) = \sup_{\alpha_1\in\Delta(A_1)}\inf_{\alpha_2\in B_\varepsilon(\alpha_1)} v_1(\alpha_1,\alpha_2,\omega) - \varepsilon$$

- Let ε-greatest commitment payoff for type ω be:

$$\overline{v}_1(\omega,\varepsilon) = \sup_{\alpha_1\in\Delta(A_1)}\sup_{\alpha_2\in B_\varepsilon(\alpha_1)} v_1(\alpha_1,\alpha_2,\omega)$$

- $\overline{v}_1(\omega, 0)$ is generalized Stackelberg payoff

Main Theorem

Theorem (3.1)**.** *For all $\varepsilon > 0$ there exists a K so that for all δ*

$$(1-\varepsilon)\delta^K\underline{v}_1(\omega_0,\varepsilon) + \left[1-(1-\varepsilon)\delta^K\right]\underline{u} \le \underline{N}_1(\delta,\omega_0)$$
$$\le \overline{N}_1(\delta,\omega_0) \le (1-\varepsilon)\delta^K\overline{v}_1(\omega_0,\varepsilon) + \left[1-(1-\varepsilon)\delta^K\right]\overline{u}.$$

- Upper bound seems weak, but is not
  - Benabou and Laroque (1988) show that a long-run player can attain utility higher than his Stackelberg payoff for low δ
  - Later we will prove that this is impossible as $\delta \to 1$
- Before proving this theorem, we state an ancillary theorem, which will be required to prove theorem 3.1

Theorem (4.1)**.** *For every $\varepsilon > 0$, $\Delta_0 > 0$ and $\Omega^+ \subset \Omega$ with $\mu(\Omega^+) > 0$ there is a $K(\varepsilon, \Delta_0, \mu(\Omega^+))$ such that for any $\mathscr{s}_1$ and σ_2, under the probability distribution generated by $\mathscr{s}_1(\Omega^+)$, there is a probability less than ε that there are more than $K(\varepsilon,\Delta_0,\mu(\Omega^+))$ periods with:*

$$\left\|p^+(h_{t-1}) - p(h_{t-1})\right\|_\infty > \Delta_0.$$

Proof of Theorem 3.1. Fix a Nash equilibrium $(\mathscr{s}_1, \sigma_2)$; $(\mathscr{s}_1, \sigma_2)$ and μ induce a joint probability distribution over types and histories.

Short-run players must use Bayesian updating in a Nash equilibrium to form posterior beliefs. Let $\alpha_2(h_{t-1})$ denote the mixed action generated by σ_2 which player 2_t plays following history h_{t-1}; similarly for $\alpha_1(h_{t-1})$ and $\alpha_1^+(h_{t-1})$. Let player 2_t's prediction of the outcome conditional on h_{t-1} and equilibrium strategies be $p(h_{t-1}) \in \Delta(Y)$. Let $p^+(h_{t-1})$ also condition on the true type being in Ω^+.

Short-run types almost have the correct distribution of outcomes even if they do not know that the long-run player's type is in Ω^+. A period is "exceptional" if short run players get a surprise in the above respect. Take $\Omega^+ = \{\omega_0\}$ and $\Delta_0 = \varepsilon$ and apply theorem 4.1. There exists a K so that in all but K periods with probability $(1-\varepsilon)$ we have:

$$\left\|p^+(h_{t-1}) - p(h_{t-1})\right\|_\infty \le \varepsilon.$$

Thus with probability $(1-\varepsilon)$ player 2_t's equilibrium action $\alpha_2(h_{t-1}) \in B_\varepsilon\left(\alpha_1^+(h_{t-1})\right)$. If player 2_t expects an outcome ε-close to $p^+(h_{t-1})$, then player 2_t must be playing a ε-confirmed best response to the mixed strategy that ω_0 would play after history h_{t-1}.

Further, since commitment types have full support, player 2_t will not play a strategy that is weakly dominated, i.e., $\alpha_2(h_{t-1}) \in B_0(\alpha_1(h_{t-1}))$.

The payoff to rational player 1 is:

$$U_1(\sigma^+,\omega_0) = E_{\sigma^+}\left[(1-\delta)\sum_{t=1}^{\infty}\delta^{t-1}v_1\left(\alpha_1^+(h_{t-1}), \alpha_2(h_{t-1}), \omega_0\right)\right].$$

Rational player's payoff in exceptional periods is bounded above by $\overline{u}$. There are at most K exceptional periods (which occur with probability greater than ε) and $U_1(\sigma^+,\omega_0)$ is maximized if these occur at the start.

Type ω_0 must want to play its equilibrium strategy and its equilibrium payoff in non-exceptional periods is at most $\overline{v}_1(\omega_0, \varepsilon)$. This proves the upper bound part of the theorem.

To prove the lower bound, use theorem 4.1 again, but take Ω^+ to be a neighborhood of the "best" commitment type for the rational long-run player.

Fix any $\alpha_1 \in \Delta(A_1)$ and take Ω^+ to be the types which play mixed strategies α_1' in the neighborhood of α_1. Let $\widetilde{\varepsilon} > 0$ be such that if $|\alpha_1' - \alpha_1|_\infty \leq \widetilde{\varepsilon}$, then $\|v_1(\alpha_1, \alpha_2, \omega_0) - v_1(\alpha_1', \alpha_2, \omega_0)\|_\infty < \varepsilon$ and $\|\rho(\cdot|(\alpha_1, \alpha_2)) - \rho(\cdot|(\alpha_1', \alpha_2))\|_\infty < \frac{\varepsilon}{2}$. Such $\widetilde{\varepsilon}$ exists since v_1 and ρ are continuous and defined on compact sets. By definition $\left|\alpha_1^+(h_{t-1}) - \alpha_1\right|_\infty \leq \widetilde{\varepsilon}$.

Apply theorem 4.1, with Ω^+ as defined above and $\Delta_0 = \frac{\varepsilon}{2}$ and note that $\mu(\Omega^+) > 0$. Suppose the rational player follows strategy α_1^+ corresponding to some commitment type in Ω^+. In non-exceptional periods, with probability at least $(1-\varepsilon)$, player 2 plays an $\frac{\varepsilon}{2}$-confirmed best responds to this strategy, but since $\|v_1(\alpha_1, \alpha_2, \omega_0) - v_1(\alpha_1', \alpha_2, \omega_0)\|_\infty < \varepsilon$, we have that in non-exceptional periods ω_0 obtains payoff at least:

$$\min_{\alpha_2 \in B_\varepsilon(\alpha_1)} v_1(\alpha_1, \alpha_2, \omega_0) - \varepsilon.$$

In exceptional periods the payoff is uniformly bounded from below by $\underline{u}$. □

Corollary (3.2)**.** *Taking the limit as $\varepsilon \to 0$ we have that:*

$$\underline{v}_1(\omega_0, 0) \leq \liminf_{\delta \to 1} \underline{N}_1(\delta, \omega_0) \leq \limsup_{\delta \to 1} \overline{N}_1(\delta, \omega_0) \leq \overline{v}_1(\omega_0, 0).$$

Proof. From Theorem (3.1) need to show that:

$$\begin{aligned} \liminf_{\varepsilon \to 1} \underline{v}_1(\omega_0, \varepsilon) &\geq \underline{v}_1(\omega_0, 0), \text{ and} \\ \limsup_{\varepsilon \to 1} \overline{v}_1(\omega_0, \varepsilon) &\leq \overline{v}_1(\omega_0, 0). \end{aligned}$$

Take $\{\varepsilon^n\}_{n=1}^\infty \to 0$ and $\alpha_2^n \in B_{\varepsilon^n}(\alpha_1)$ for all n and note that $\lim_n \alpha_2^n \in B_0(\alpha_1)$. □

- A game is non-degenerate if $\nexists a_2 \in A_2$ which is undominated such that for some $\alpha_2 \neq a_2$, $v(\cdot, \alpha_2) = v(\cdot, a_2)$.
  - Satisfied for an open, dense set of payoffs
- A game is identified if for each α_2 that is not weakly dominated $\rho(\cdot|\alpha_1, \alpha_2) = \rho(\cdot|\alpha_1', \alpha_2)$ implies $\alpha_1 = \alpha_1'$.

Theorem (3.3)**.** *In a non-degenerate, identified game $\underline{v}_1(\omega_0, 0) = \overline{v}_1(\omega_0, 0)$.*

- Generically, average payoff of a patient long-run player in any NE is determined by reputation effects if actions are observed

Proof of Theorem 3.3. Since the game is identified $B_0(\alpha_1)$ the set of 0-confirmed best responses is simply the set of undominated best responses to α_1. Suffices to show that for $\alpha_2 \in B_0(\alpha_1)$, there exists a sequence $\{\alpha_1^n\}_{n=1}^\infty$ which converges to α_1 such that:

$$\{\alpha_2\} = B_0(\alpha_1^n).$$

There exists some mixed action $\alpha_1' \in \Delta(A_1)$ such that α_2 is a strict best response to α_1'. Take a sequence $\{\kappa^n\}_{n=1}^\infty$ such that $\kappa^n \in (0,1)$ and $\kappa^n \to 1$. Define $\alpha_1^n = \kappa^n \alpha_1 + (1-\kappa^n)\alpha_1'$. Note that α_2 is a strict best response to α_1^n. □

Remarks about the Technical Result

- The main technical contribution of the paper is theorem 4.1, restated here for convenience

Theorem (4.1)**.** *For every $\varepsilon > 0$, $\Delta_0 > 0$ and $\Omega^+ \subset \Omega$ with $\mu(\Omega^+) > 0$ there is a $K(\varepsilon, \Delta_0, \mu(\Omega^+))$ such that for any $\mathcal{S}_1$ and σ_2, under the probability distribution generated by $\mathcal{S}_1(\Omega^+)$, there is a probability less than ε that there are more than $K(\varepsilon, \Delta_0, \mu(\Omega^+))$ periods with:*

$$\left\|p^+(h_{t-1}) - p(h_{t-1})\right\| > \Delta_0.$$

- To prove the above, first show that the odds ratio is a supermartingale (lemma 4.1)
  - Supermartingales converge almost surely, but not uniformly
  - Fudenberg and Levine show that *active* supermartingales converge uniformly
- To show the rest of the theorem, note that in exceptional periods, there is a substantial (i.e., greater than Δ_0) probability that the short run player will be substantially wrong in their forecast
  - Thus, the supermartingale L_t is active, in the sense that L_t has a significant probability of decreasing by a sizable fraction
  - Use the level of activity of a supermartingale to get a bound for the number of exceptional periods
- Sorin (1999) remarks that Theorem 4.1 is a "uniform version" of the merging of beliefs theorem by Blackwell and Dubins (1962)
  - Blackwell and Dubins (1962) consider when posterior beliefs of individuals will merge, if individuals start with different priors and observe the same outcomes

Concluding Remarks

- Introducing reputation yields a sharp prediction for the payoff of patient long-run players
- Generically, if the long-run player's action is statistically identified, the long-run player obtains his Stackelberg payoff