Lecture 9: Reputation Effects in Repeated Games

Alexander Wolitzky

MIT

14.126, Spring 2024

Reputation in Repeated Games

Previous 4 lectures considered repeated game models of **cooperation**: when and how the prospect of future rewards/punishments motivates non-static Nash behavior.

- Simplest model has multiple long-run players + complete info.

Next 2 lectures consider repeated game models of **reputation formation**: when and how the desire to signal something about your type drives dynamic behavior.

- Simplest model has 1 long-run player facing series of short-run opponents + incomplete info about LR player's types.
- The LR player's **reputation** at a given history is her opponents' belief about her type.

Reputation in Repeated Games (cntd.)

Formally, reputation models are a special class of repeated games with incomplete information.

In general, repeated games with incomplete info can be complicated, because there are many different kinds of strategic consideration: signaling, revealing/concealing information, learning and experimentation, etc..

- Successful analysis tends focuses on one issue at a time.

"Reputation model" generally means (albeit with exceptions):

- Only one player with private info.
- The private info is whether the player is **rational** (maximize expected discounted utility) or a **commitment type** (who takes a fixed strategy, i.e. has a strictly dominant repeated game strategy).
- The player with private info is a long-run player with $\delta \to 1$, facing a series of short-run opponents.

Why Study Reputation in Repeated Games?

While reputation models are "special" in various ways, they're important for a couple reasons.

1. Useful for studying long-run implications of signaling in **applications**.

- How do firms build reputations for providing high quality?
- When do firms with good reputations maintain them vs. run them down?
- How do brand names or labels indicate reputation?
- When/how does desire for a good reputation lead to perverse "pandering"-type incentives?

Why Study Reputation in Repeated Games? (cntd.)

2. Reputation models give sharp results on **equilibrium selection** in repeated/dynamic games.

We've seen that complete info models often yield a folk theorem.

In contrast, reputation models often make sharp payoff predictions.

Two leading examples of this are:

- The **Stackelberg payoff theorem**: in canonical LR-SR reputation models, the LR player gets her Stackelberg payoff as $\delta \to 1$. Idea: if LR player always takes the Stackelberg action, SR players eventually come to expect this, and hence best respond. (Today's lecture.)
- The **reputational Coase conjecture:** in bilateral bargaining with one possibly tough bargainer, that player does as well as if known to be tough for sure. (Next week.)

Plan

1. Early reputation results (KWMR 82, FM 86).
2. The Stackelberg payoff theorem and extensions (FL 89, 92, Gossner 11)
3. Long-run implications of reputation (Cripps-Mailath-Samuelson 04)
4. Extensions (interdependent values, long-run "audience," multiple reputation-builders)

Next lecture: reputation effects in markets, covering both "good reputation" (Holmström 82, Mailath-Samuelson 01, Board-Meyer-ter-Ven 11) and "bad reputation" (Ely-Valimäki 03, Ely-Fudenberg-Levine 08).

The Chain Store Paradox

Start with a canonical example (and inspiration for seminal papers of Kreps-Wilson-Milgrom-Roberts 82): the chain store game.

- A LR player ("incumbent") faces a series of T SR players ("entrants").
- In each period t, the current entrant chooses *Enter* or *Stay Out*. If enters, the incumbent chooses *Fight* or *Accommodate*. These decisions are observed by future entrants.
- Stage game payoffs are

	Enter	*Stay Out*	
Fight	$-1, -1$	$a, 0$	$(a, b > 0)$
Accomodate	$0, b$	$a, 0$	

- Incumbent maximizes undiscounted sum of payoffs.
- Assume $a > 1$, so incumbent willing to fight in period t to deter entry in period $t + 1$.

Chain Store Paradox (cntd.)

With complete info, by backward induction the game has a unique SPE: all entrants always enter, and incumbent always accommodates.

- This is the classic **chain store paradox** (Selten 78).

It's called a paradox because it seems unreasonable. If $T = 1000$ and the first 100 entrants entered and got fought, what should entrant 101 do? (The backward induction prediction just got rejected 100 times over...)

KWMR propose a natural resolution (anticipated by Nash): Suppose that with small prob $\mu_0 > 0$ ("initial reputation"), the incumbent is a "tough type" that always fights.

- For any $\mu_0 > 0$, not a SE for play to procede as in the complete info game when incumbent is rational, because playing *Fight* in period 1 would jump posterior reputation μ up to 1 and deter all future entry. So what happens instead?

The Gang of Four Theorem

Theorem

For any $\mu_0 > 0$, there is a number T^*, independent of T, s.t. in any SE the entrants stay out until the last T^* periods. Hence, as $T \to \infty$, the incumbent's average payoff conveges to a.

- The discontinuity at μ_0 resolves the chain store paradox.
- Proof proceeds by backward induction, uniquely characterizing eqm behavior in each period t as a function of the incumbent's reputation at the beginning of the period μ_t.
- Shows that there exists a number $\pi \in (0, 1)$ (independent of T) s.t. entrant stays out in period t (and incumbent fights in period t with prob 1 if entrant enters) if $\mu_t > \pi^{T-t}$.
- When $\mu_0 > \pi^T$, this implies that entrants stay out in "early periods," and hence reputation remains at μ_0.
- "Early periods" last until final $\log \mu_0 / \log \pi$ periods, which is a constant independent of T.

Remarks

KWMR also showed that that the same idea resolves other paradoxes of backward induction.

- In the finitely repeated PD, players cooperate until the last few periods, if there is a small prob that each player is committed to Tit-for-Tat.
- In the centipede game, players *Pass* until the last few periods, if there is a small prob that each player is committed to *Always Pass*.

It's natural that players think opponents could have very different preferences/strategies with small prob. However, an apparent limitation of KWMR is **prior-dependence:** the prediction seems to depend on the particular commitment type we threw in.

- E.g. a small chance the incumbent is committed to *Always Accommodate* doesn't change anything.
- This point is made more generally by FM 86's "folk theorem with incomplete information" (albeit for simultaneous-move games with all LR players).

FM 86

Theorem

Fix a (simultaneous-move) stage game and a static NE payoff vector v. For any $\varepsilon > 0$ and any payoff vector $v' > v$, there exists $\bar{T}$ such that, for any $T > \bar{T}$, there exists a strategy s_i for each player i in the T-period finitely repeated game where each player i is rational with prob $1 - \varepsilon$ and is committed to s_i with prob ε s.t. there is a SE where average payoffs are within ε of v'.

- Let s_i be the "trigger strategy" that targets v', switches to static NE after any deviation.
- Suppose rational players also play s_i until the last $\bar{T}$ periods, for some fixed large $\bar{T}$.
- Then deviating from s_i before the last $\bar{T}$ periods is unprofitable, as if deviate get continuation payoff v_i, if conform get continuation payoff at least $\varepsilon^{N-1} v_i' + \left(1 - \varepsilon^{N-1}\right) v_i$.
- And deviations in the last $\bar{T}$ periods have little effect on average payoffs.

Toward the Stackelberg Payoff Theorem

FM 86 shows that if tailor commitment type to target payoff, can attain a wide range of payoffs.

- Does this preclude a robust reputation result?
- Not if the "tailored" type spaces are themselves special, and we actually get sharp predictions for "most" type spaces: e.g., those with a wide range of commitment types.
- But, a challenge: how to characterize equilibria with many commitment types?

Toward the Stackelberg Payoff Theorem (cntd.)

A key of insight of FL 89, 92: Don't exactly characterize eqm strategies, instead find bounds on eqm payoffs.

While eqm strategies may be prior-dependent, we'll see that in LR-SR game as $\delta \rightarrow 1$ can often make sharp, prior-independent payoff predictions.

Idea: if LR player always takes the Stackelberg action, SR players eventually come to expect this, and hence best respond.

- This logic works even if the prior puts weight on many commitment types, so long as it includes the Stackelberg commitment type.
- The Stackelberg type is thus **canonical**.

Reputation with Perfect Monitoring (FL 89)

A LR player (player 1) with discount factor δ faces an infinite sequence of SR opponents (player 2's). Finite stage game (A, u).

LR player's type drawn from prior μ that can put weight on the **rational** type of player 1 and also **commitment** types θ_{a_1} for each $a_1 \in A_1$. Commitment type θ_{a_1} always takes a_1.

The **Stackelberg action** a_1^* is the pure action player 1 would most like to publicly commit herself to, assuming player 2 takes his worst best response:

$$a_1^* \in \operatorname*{argmax}_{a_1} \min_{a_2 \in BR_2(a_1)} u_1(a_1, a_2).$$

Let $u_1^* = \min_{a_2 \in BR_2(a_1^*)} u_1(a_1^*, a_2)$.

- E.g., in the chain store game, $a_1^* = \mathit{Fight}$, $u_1^* = a$.

Theorem

Fix any prior μ s.t. $\mu(\theta_{a_1^*}) > 0$. For all $\varepsilon > 0$, there exists $\bar{\delta} < 1$ such that if $\delta > \bar{\delta}$ then in any NE rational player 1's payoff is at least $u_1^* - \varepsilon$.

Intuition

Suppose player 1 always takes a_1^*.

- If in some period t player 2 puts low prob on a_1^* being played, then if a_1^* is played, by Bayes' rule player 1's reputation ($=\mu_t\left(\theta_{a_1^*}|h^t\right)$) jumps up by a multiple bounded away from 1.
- $\mu_t\left(\theta_{a_1^*}\right)$ is bounded by 1, so it can only jump up finitely many times (indep of δ).
- So, can be only finitely many periods where player 2 puts low prob on a_1^*. These are the only periods where player 2 can fail to best respond to a_1^*.
- So player 1 gets at least $u_1^* - \varepsilon$.

Since player 1 gets at least $u_1^* - \varepsilon$ if always takes a_1^*, must also get at least this much in eqm.

- Does **not** imply that player 1 always takes a_1^* in every eqm.

Remarks

- Striking **equilibrium selection** result. The folk theorem says "anything goes" as $\delta \to 1$. Here, precise payoff prediction when $\delta \to 1$. Higher δ pins prediction down **more precisely**.
- Prior can put arbitrarily small weight on $\theta_{a_1^*}$. There is an order-of-limits issue between $\mu\left(\theta_{a_1^*}\right)$ and δ (for any δ, the payoff bound vanishes as $\mu\left(\theta_{a_1^*}\right) \to 0$), but the proof shows that ε can be taken proportional to $(1-\delta)\log\mu\left(\theta_{a_1^*}\right)$, so OK for $\mu\left(\theta_{a_1^*}\right)$ to be much smaller than $1-\delta$.
- Prior can also put weight on other commitment types.
- Theorem holds for any **Nash** eqm. (Actually, holds whenever SR players take a best response to **some** strategy for rational LR player, and rational LR player takes a best response to that. That is, 2 rounds of deletion suffices. See Watson 93, Battigalli-Watson 97.)
- Always taking a_1^* does **not** necessarily convince SR players that LR player is **committed** to a_1^*. Just convinces them that she will **play** a_1^*. This is enough to guarantee payoff u_1^*.

Proof

Lemma

In any NE, for any $q < 1$, along a history where LR always takes a_1^*, there are at most $\log \mu_0 \left(\theta_{a_1^*}\right) / \log q$ periods where $\Pr\left(a_1^* | h^t\right) < q$.

- By Bayes' rule,

$$\mu_{t+1}\left(\theta_{a_1^*} | h^t, a_1^*\right) = \frac{\mu_t\left(\theta_{a_1^*} | h^t\right)}{\Pr\left(a_1^* | h^t\right)}.$$

- So, when LR always takes a_1^*, the sequence $\left(\mu_t\left(\theta_{a_1^*} | h^t\right)\right)_t$ is non-decreasing, and it increases by a multiple of at least $1/q$ whenever $\Pr\left(a_1^* | h^t\right) < q$.
- Since $\mu_t\left(\theta_{a_1^*} | h^t\right) \leq 1$, the number of periods with $\Pr\left(a_1^* | h^t\right) < q$ can't exceed any T s.t.

$$\frac{\mu_0\left(\theta_{a_1^*}\right)}{q^T} \geq 1, \text{ or } T \leq \frac{\log \mu_0\left(\theta_{a_1^*}\right)}{\log q}.$$

Proof (cntd.)

Since A_2 is finite, $\exists q < 1$ s.t. every best response when $\Pr(a_1^* | h^t) \geq q$ is a best response to a_1^*.

So, if LR always takes a_1^*, there are at most $K := \log \mu_0 (\theta_{a_1^*}) / \log (q)$ periods where SR does not take a best response.

- Hence, LR gets at least u_1^* in all but at most K periods.
- LR's payoff from *Always* a_1^* is as least $\left(1 - \delta^{K+1}\right) \underline{u}_1 + \delta^{K+1} u_1^*$.
- This $\rightarrow u_1^*$ as $\delta \rightarrow 1$.
- Hence, LR's eqm payoff must also $\rightarrow u_1^*$ as $\delta \rightarrow 1$.

Reputation with Imperfect Monitoring (FL 92)

Now suppose LR player's action is imperfectly observed: SR players observe $y \sim p\left(\cdot|a_1\right)$.

This more general model also covers commitment to mixed actions (as the realized action is a signal of the mixed action) and extensive-form stage games (as the realized outcome is a signal of the stage-game strategy).

To cover extensive-form stage games, we won't assume LR player's action is identified.

Instead, define a generalization of the best response correspondence, which allows SR players to have incorrect beliefs about LR player's play at unreached subgames.

Epsilon-Confirmed Best Responses

Definition

$\alpha_2 \in \Delta(A_2)$ is an ε-**confirmed best response to** $\alpha_1 \in \Delta(A_1)$ if it's not weakly dominated and there exists α_1' s.t.

1. $\alpha_2 \in \operatorname{argmax}_{\hat{\alpha}_2} u_2(\alpha_1', \hat{\alpha}_2)$.
2. $|p(y|\alpha_1, \alpha_2) - p(y|\alpha_1', \alpha_2)| < \varepsilon \ \forall y$.

Intuitively, "ε-confirmed BR to α_1" means BR to some action that's almost indistinguishable from α_1.

Let $BR_2^{\varepsilon}(\alpha_1)$ denote the set of ε-confirmed best responses to α_1.

Stackelberg Payoff Theorem

Theorem

Fix any prior μ and any $\alpha_1 \in \Delta(A_1)$ s.t. $\mu(\theta_{\alpha_1}) > 0$. For all $\varepsilon > 0$, there exists $\bar{\delta} < 1$ such that if $\delta > \bar{\delta}$ then in any NE rational player 1's payoff is at least

$$\min_{\alpha_2 \in BR_2^{\varepsilon}(\alpha_1)} u_1(\alpha_1, \alpha_2) - \varepsilon.$$

Intuition: if LR always takes α_1^*, with high prob there are only finitely many periods where SR expects signals far from those under α_1^*; so SR almost always best responds to a strategy that generates signals close to those under α_1^*.

Proof Approaches

FL 92 prove this using martingale techniques.

- If LR always takes α_1^*, $\mu_t\left(\theta_{\alpha_1^*}|h^t\right)$ no longer goes up deterministically, but it's a submartingale.
- Can use martingale arguments to show that for any $q < 1$, there exists K independent of δ s.t. with high prob $\Pr\left(\left\{\alpha_1 : \left\|p\left(\cdot|\alpha_1, \alpha_2\left(h^t\right)\right) - p\left(\cdot|\alpha_1^*, \alpha_2\left(h^t\right)\right)\right\| < \varepsilon\right\}|h^t\right) < q$ for at most K periods.
- Sorin 99 gives a shorter proof using results on "merging" of Bayesian posteriors. (See MS Ch. 15.4.)

Gossner 11 gives similar (in some cases stronger) results with short proofs using entropy methods.

- Idea is to bound SR players' "expected prediction errors" over T periods.

FL, Sorin, and Gossner use different probabilistic methods to formalize similar basic idea.

Reputation in the Long Run

The Stackelberg payoff theorem concerns players' expected payoffs from the beginning of the game.

In some settings, we might also be interested in long-run ($t \to \infty$) behavior.

- Average welfare of all SR players (or long-run welfare in overlapping generations models more generally).
- Steady state predictions (especially in models without a well-defined start date).

At first one might think that if LR player gets average payoff $\to u_1^*$ as $\delta \to 1$, must get long-run payoff close to u_1^*.

But this confuses the order of limits: average payoff is determined by first $O\left(1/\left(1-\delta\right)\right)$ periods, so for any δ long-run payoffs could be far from u_1^*.

- Somewhat surprisingly, under some conditions this is actually what happens.

Cripps-Mailath-Samuelson 04

LR vs. SR model with imperfect monitoring. Assumptions:

- Full support: $p(y|a_1) > 0 \; \forall y, a_1$.
- Player 1's action is identified: $\forall \alpha_2$, $\alpha_1 \neq \alpha_1' \implies p(\cdot|\alpha_1, \alpha_2) \neq p(\cdot|\alpha_1', \alpha_2)$.
- Countable set of commitment types. Player 2 has a unique BR to each commitment type's action.

Theorem

Let $\hat{\Theta}$ be the set of types θ_{a_1} s.t. $(a_1, BR_2(a_1))$ is not a static NE. In any NE, when LR is rational, $\lim_{t\to\infty} \mu\left(\hat{\Theta}|h^t\right) = 0$ almost surely.

- FL: LR payoff$\rightarrow u^*$ as $\delta \rightarrow 1$.
  CMS: for any δ, LR reputation$\rightarrow 0$ as $t \rightarrow \infty$.

Intuition

- Suppose LR takes a_1^* for a long time, so SR expects a_1^* for a long time with high prob.
- Due to full support, eventually SR will continue to expect a_1^* for a long time regardless of current-period signal.
- Since $\left(a_1^*, BR_2\left(a_1^*\right)\right)$ is not a static NE, LR will deviate at such a history.
- Thus, while LR can guarantee a payoff close u^* by always taking a_1^* (by FL), she does even better by occasionally deviating from a_1^*.
- These deviations may be infrequent, but they occur infinitely often.
- Repeated deviations eventually run down LR's reputation.

Recovering Permanent Reputation

The CMS "impermanent reputation" result is a bit surprising, because we do seem to see reputation effects even in interactions that have been going on for a long time.

The literature has considered a couple ways of recovering long-run reputation effects.

1. Changing types. Ekmekci-Gossner-Wilson 12 show that with a constant, iid prob ρ that LR's type is redrawn from the prior, LR's expected payoff starting at any on-path history obeys the same bound as the ex ante payoff with initial reputation $\rho\mu\left(\theta_{a^*}\right)$.

- This bound goes to 0 as $\rho \to 0$ but convergence is very slow in ρ, since the payoff bound is approximately $u_1^* - (1-\delta)\log\left(\rho\mu\left(\theta_{a^*}\right)\right)$. So even a small prob of type changes restores long-run reputation.

Recovering Permanent Reputation

2. Bounded memory. Long-run reputation effects also arise if it's the SR players' information that turns over, rather than the LR player's type. Liu-Skrzypacz 14 consider a model along these lines.

- Consider a class of "trust games" where SR chooses how much to trust LR, LR chooses how much to exploit SR. Assume LR's gain from exploiting is greater when SR is more trusting. Stackelberg action is *don't exploit*.
- In eqm, LR either doesn't exploit or exploits maximally. SR's level of trust depends only on number of periods since the most recent exploitation, increasing gradually over time to keep LR indifferent.
- At "clean" histories where LR has not exploited in the last K periods, LR always exploits.
- Thus, get a theory of infinitely recurring reputation cycles.
- (Earlier papers by Sobel 85, Benabou-Laroque 92 had somewhat similar dynamics.)

Extensions of the Baseline Reputation Model

There are many. We'll discuss a few:

1. Interdependent values: SR directly cares about LR's type, not just her action.
2. Long-run player 2 (with known type).
3. Two LR reputation builders.

Interdependent Values (Pei 20)

Suppose there are multiple rational types θ (in addition to commitment types), and SR's best response $BR_2(\alpha_1, \theta)$ depends on both α_1 and θ. (Assume it's single-valued.)

- E.g., SR players want to eat at LR player's restaurant iff LR player is **both** a good cook (high θ) and works hard (high a_1). ("Product choice game with unknown quality.")

Natural definition of Stackelberg payoff for rational type θ here is

$$u_1^*(\theta) = \max_{\alpha_1} u_1(\alpha_1, BR_2(\alpha_1, \theta); \theta).$$

Question: When is each rational type θ guaranteed a payoff close to $u_1^*(\theta)$ in every NE when $\delta \to 1$?

Challenge: If type θ takes a_1 repeatedly, this does convince SR that she'll keep taking a_1, just like in FL 89. However, it need **not** convince SR that LR's type is θ, and inducing $BR_2(\alpha_1, \theta')$ for $\theta' \neq \theta$ may be bad for type θ. There might be a tradeoff between establishing a reputation for taking a_1 and signaling favorable information about θ.

Interdependent Values (cntd.)

Pei shows that, in general, rational types are not guaranteed a payoff close to $u_1^*(\theta)$ in every NE as $\delta \to 1$.

However, this does hold if $|A_2| = 2$ and the game is "monotone-supermodular": u_1 is decreasing in a_1 and increasing in a_2; u_1 has strictly increasing differences in θ and (a_1, a_2); u_2 has strictly increasing differences in a_2 and (θ, a_1).

Intuition: for this class of games, higher LR actions convince SR both that θ is high and that LR will keep taking higher actions.

For example, these conditions are satisfied in the product choice game with unknown quality.

Long-Run Player 2

What happens if the “audience” for reputation-building is a LR player instead of a series of SR players?

- E.g., worker trying to prove herself to the market (myopic “SR” player, which sets wage=productivity each period) vs. a boss (strategic LR player).

If players 1 and 2 are equally patient, there’s a profound problem: # of times P1 is willing to take a_1^* to convince P2 to BR is proportional to how convinced P2 must become before he must start taking a BR. (More on this in a moment.)

More promising case (also closer to the baseline model): P2 is patient, but P1 is much more patient.

Formally, is the iterated limited payoff $\lim_{\delta_2 \to 1} \lim_{\delta_1 \to 1} u_1$ always close to the Stackelberg payoff u_1^*?

- This is sometimes called the “long-run/medium-run model.”

Long-Run Player 2 (cntd.)

Turns out that there is also a challenge here:

- If P1 takes a_1^* repeatedly, this eventually convinces P2 that she will keep taking a_1^* for a long time with high prob **on path**, i.e., with respect to the eqm prob dist. (Just like in FL 89.)
- However, if P2 is not taking a BR in eqm, this does **not** convince P2 that P1 will keep taking a_1^* if P2 switches to taking a BR.
- (This **is** what would happen if P1 were the Stackelberg **type**. But P1 can only convince P2 that she will take the Stackelberg **action**.)
- E.g., P2 can suspect that P1 is playing the strategy "Take a_1^* until P2 takes a BR, then switch to minmaxing P2." Then P2 should not BR, no matter how many times P1 takes a_1^*.

For this reason, FL 89 theorem does not hold in the long-run/medium-run case.

Long-Run Player 2 (cntd.)

However, there are several variations of the model that do yield reputation effects in the long-run/medium-run case.

1. P1 can guarantee the "minmaxing Stackelberg payoff"

$$\max_{a_1 : u_2(a_1, BR_2(a_1)) = \underline{u}_2} \; \min_{a_2 \in BR_2(a_1)} u_1(a_1, a_2),$$

because if a_1 already minmaxes P2, P2 has nothing to fear from a switch and so will BR. (Schmidt 93, extended by Cripps-Schmidt-Thomas 96)

2. P1 can guarantee the Stackelberg payoff u_1^* under full-support public monitoring. Intuitively, full support noise blurs the distinction between on-path and off-path histories, so P2 learns P1's strategy, not just her on-path action. More precisely, divide the game into T-period blocks. If P1 repeatedly plays the same T-period repeated game strategy, P2 eventually learns this strategy and best responds (if it has positive ex ante prob). (Aoyagi 96, Celentani-Fudenberg-Levine-Pesendorfer 96)

Long-Run Player 2 (cntd.)

3. If P1 could be a commitment type that punishes P2 for **not** taking a BR, then eventually P2 should BR when P1 plays like this type. (Evans-Thomas 97)

Two LR Reputation-Builders

Finally, consider "symmetric" situations where players are equally patient and there's incomplete info on both sides.

Typically, the players' desired commitments conflict, so it's impossible for each to get her Stackelberg payoff. E.g., this happens in the battle-of-the-sexes:

	B	O
B	$2, 1$	$0, 0$
O	$0, 0$	$1, 2$

Each player's Stackelberg payoff is 2, but they can't both get 2.

- A natural guess is that in this type of game, eqm play will resemble a war of attrition, where both players try to establish a reputation, and eventually one of them admits that they're rational and best responds.
- Next week, we'll see that this is what happens in "reputational bargaining" models, which are related to the battle-of-the-sexes with 2-sided reputation.

2-Sided Reputation (cntd.)

However, it turns out that with 2 LR players, the Stackelberg payoff theorem can fail even when the players' desired commitments agree!

Consider

	B	O
B	$1, 1$	$0, 0$
O	$0, 0$	$0, 0$

where each player is committed to B with prob μ, rational otherwise.

We might think that by repeatedly taking B they can force coordination on B.

But:

Theorem (Cripps-Thomas 97)

For any $\varepsilon > 0$, there exist $\bar{\mu} > 0$ and $\bar{\delta} < 1$ s.t., for all $\mu < \bar{\mu}$ and $\delta > \bar{\delta}$, there exists a SE where both players' payoffs are less than ε.

2-Sided Reputation (cntd.)

Problem: # times P1 is willing to take a_1^* to convince P2 to BR is proportional to how convinced P2 must become before he must start taking a BR.

We would get a reputation result here in the long-run/medium-run case, or with full-support imperfect monitoring (Cripps-Faingold 14), or if the players' strategies have bounded recall (Aumann-Sorin 89).

In general, games with 2 LR reputation-builders are subtle and not well-understood.

- Atakan-Ekmekci 12, 15 give conditions under which one LR reputation-builder can secure Stackelberg payoff against an equally patient LR opponent.
- Best-studied class of games with 2 LR reputation-builders is reputational bargaining, which we'll cover next week.