Lecture 8: Topics in Repeated Games

Alexander Wolitzky

MIT

14.126, Spring 2024

Topics in Repeated Games

Last 3 lectures: foundational papers on repeated games with imperfect public monitoring by APS and FLM.

Today: quick tour of three major further topics in repeated games.

- **Monitoring vs. discounting:** folk theorem says we get cooperation when $\delta \to 1$ for fixed monitoring and stage game, but what can be said about how discounting, monitoring, and the stage game jointly determine prospects for cooperation?
- **Private monitoring:** how does private monitoring make cooperation harder and what can be done about it? Can private monitoring ever make cooperation easier?
- **Community enforcement:** can a group support cooperation when people play with different partners each period? How does this relate to standard repeated games with fixed partners?

Cover some theorems but mostly big-picture guide to the literature.

Monitoring vs. Discounting: Examples

Frequent actions. (Abreu Milgrom Pearce 1991; Fudenberg Levine 2007; Sannikov Skrzypacz 2007)

- Suppose actions affect an underlying continuous-time signal process, and period length Δ measures how frequently players observe the process and potentially change actions.
- E.g., sales are essentially continuous, but management observes them and adjusts strategy every Δ days.
- Then $\Delta \to 0$ implies little discounting between updates, but also little information.
- Sannikov (2007) develops a continuous-time repeated game model where actions determine the drift of a publicly observed Brownian motion. Can interpret as a natural limiting case as $\Delta \to 0$. Characterizes boundary of PPE payoff set as solution to an ODE.

Monitoring vs. Discounting: Examples (cntd.)

Large populations. (Fudenberg Levine Pesendorfer 1998; al-Najjar Smorodinsky 2000; Sugaya Wolitzky 2024)

- Suppose players are patient ($\delta \to 1$) but there are many of them ($(1-\delta)\,N \nrightarrow 0$), and they are monitored through some "aggregate signal."
- What properties of the monitoring structure determine the prospects for cooperation in this case?

Monitoring vs. Discounting: General Tradeoff

Sugaya Wolitzky 2023 derive general results on the tradeoff between monitoring and discounting of the following form: there is a measure of monitoring precision χ^2 such that if $\chi^2/(1-\delta)$ is small then cooperation is impossible (play is "ε-myopic"), and if $\chi^2/(1-\delta)$ is large then cooperation is possible (a folk theorem holds).

- Impossibility result holds for public or private monitoring; folk theorem is proved for public monitoring (as we'll see, proving folk theorems with private monitoring is hard).
- The boundary case where $\chi^2 \to 0$ and $\delta \to 1$ at the same rate is the frequent action limit. In this case, typically partial cooperation is possible, e.g. as characterized in the "continuous-time limit" by Sannikov.

We cover the impossibility result (more novel); skip the folk theorem (builds on FLM).

The Blind Game

First observation: by a "revelation principle," for any monitoring structure (Y, p), the set of NE outcomes is smaller than that in the following **blind game.**

- There is a mediator. At the beginning of the each period, mediator privately recommends an action to each player.
- At end of each period, players observe **nothing** (except own action); mediation observes y (drawn with prob $p(y|a)$).
  - Mediator does **not** observe a.
- Player i's period-t action can depend on $\left((a_{i,t'}, r_{i,t'})_{t'=1}^{t-1}, r_{i,t}\right)$, where $r_{i,t}$ is mediator's period-t' recommendation to player i.

Intuition: blind game has same "amount of information" as the original game, but the information is optimally distributed.

Outcomes and Occupation Measures

An **outcome** of the repeated game $\mu \in \Delta\left((A \times Y)^{\infty}\right)$ is a distribution over infinite paths of action profiles and signals.

- A strategy profile induces a unique outcome.

The outcome determines the ex ante marginal distribution over period-t action profiles, $\alpha_t^{\mu} \in \Delta(A)$.

The **occupation measure over actions** induced by μ, $\alpha^{\mu} \in \Delta(A)$, is the expected discounted fraction of periods each action profile is played:

$$\alpha^{\mu} = (1-\delta)\sum_{t=1}^{\infty} \delta^{t} \alpha_t^{\mu}.$$

Note: occupation measure determines payoffs.

- Expected payoff vector under outcome μ is $u(\alpha^{\mu})$.
- α^{μ} is how the game is played "on average."

Manipulations

In a mediated game, a 1-shot deviation is a **manipulation** $s_i : A_i \to \Delta(A_i)$.

- When recommended a_i, play $s_i(a_i)$ instead.

The **gain** from manipulation s_i at action profile distribution $\alpha \in \Delta(A)$ is

$$g_i(s_i, \alpha) = \sum_{a \in A} \alpha(s)\left(u_i\left(s_i(a_i), a_{-i}\right) - u_i(a)\right).$$

For any $\varepsilon > 0$, an action profile distribution $\alpha \in \Delta(A)$ is a **static** ε-**correlated equilibrium** if $g_i(s_i, \alpha) \leq \varepsilon$ for all i and s_i.

Detectability

The **detectability** of manipulation s_i at action profile a is

$$\chi_i^2(s_i, a) = \sum_y p(y|a) \left(\frac{p(y|s_i(a_i), a_{-i}) - p(y|a)}{p(y|a)} \right)^2 .$$

- χ^2-divergence of $p(\cdot|s_i(a_i), a_{-i})$ from $p(y|a)$.
- Well-defined if signals have full support (assume this).
- Extend linearly to $\chi_i^2(s_i, \alpha)$ for $\alpha \in \Delta(A)$.

Intuition:

- Likelihood ratio difference is key info measure for incentives (Mirrlees 1975, Holmstrom 1979).
- Expected likelihood ratio difference is always 0.
- Likelihood ratio difference is "often large" iff its variance is large.
- χ^2-divergence is the variance of the likelihood ratio difference.

Why is \chi^2-Divergence the "Right Measure"?

Punishment for deviating=change in E[continuation payoff]:

$$\begin{aligned} & \sum_y \left(p\left(y|s_i\left(a_i\right), a_{-i}\right) - p\left(y|a\right)\right) w_i\left(y\right) \\ = \quad & \sum_y p\left(y|a\right)\left(\frac{p\left(y|s_i\left(a_i\right), a_{-i}\right) - p\left(y|a\right)}{p\left(y|a\right)}\right)\left(w_i\left(y\right) - \mathbb{E}\left[w_i\left(y\right)\right]\right). \end{aligned}$$

By Cauchy-Schwarz for the inner product $\langle X, Y\rangle = \sum_y p\left(y|a\right) X\left(y\right) Y\left(y\right)$, an upper bound on the punishment is

$$\sqrt{\chi_i^2\left(s_i, a\right) \mathrm{Var}\left(w_i\left(y\right)\right)}.$$

To deter deviations, χ^2-divergence and continuation payoff variance must both be large.

Bound is tight when likelihood ratio differences and continuation payoff differences are co-linear.

- Can show this is optimal for minimizing continuation payoff variance under incentive constraints.

Theorem: Mnemonic

Theorem

For any NE in any repeated game,

$$\textit{deviation gain} \leq \sqrt{\frac{\delta}{1-\delta}\textit{(detectability)(on-path payoff variance)}}.$$

Theorem: Formal

Theorem

For any NE in any repeated game (as well as in the corresponding blind game), any player i, and any manipulation s_i, we have

$$g_i\left(s_i, \alpha^{\mu}\right) \leq \sqrt{\frac{\delta}{1-\delta}\chi_i^2\left(s_i, \alpha^{\mu}\right) V_i\left(\alpha^{\mu}\right)}.$$

In particular, α^{μ} is a static ε-correlated equilibrium (and hence repeated game payoffs under μ are static ε-correlated eqm payoffs), for

$$\varepsilon = \max_{i, s_i} \sqrt{\frac{\delta}{1-\delta}\chi_i^2\left(s_i, \alpha^{\mu}\right) V_i\left(\alpha^{\mu}\right)}.$$

Remarks

$$\varepsilon = \max_{i,s_i} \sqrt{\frac{\delta}{1-\delta} \chi_i^2 \left(s_i, \alpha^\mu\right) V_i \left(\alpha^\mu\right)}.$$

- Incentives bounded by square root of product of patience, detectability, and on-path payoff variance.
- FLM folk theorem limit: small detectability OK when $\delta \to 1$.
- Perfect monitoring limit: on-path variance $\to 0$ OK when detectability $\to \infty$.
- "Surprising" that $(1-\delta)^{-1}$ goes inside square root, as continuation payoffs get weight $(1-\delta)^{-1}$ relative to stage-game payoffs.
- However, impossible for all entire continuation payoff to depend fully on each current period signal.

Intuition for 1-\delta Order

Why is $(1-\delta)^{-1/2}$ the right order?

- Consider "review strategy" that aggregates information for T periods before deciding on reward/punishment.
  - Not exactly optimal but suffices for folk theorem: Radner 1985.

Intuition for 1-\delta Order

Why is $(1-\delta)^{-1/2}$ the right order?

- Consider "review strategy" that aggregates information for T periods before deciding on reward/punishment.
  - Not exactly optimal but suffices for folk theorem: Radner 1985.
- Longest possible review is $T \approx \left((1-\delta)^{-1}\right)$. With this review length, present value of rewards/punishments is independent of δ.
- Standard deviation of signal counts is at least $(1-\delta)^{-1/2}$.
- Probability that a 1-shot deviation is pivotal is at most $(1-\delta)^{1/2}$.
- Hence, average incentive strength is at most $(1-\delta)^{1/2}$.
- Average static deviation gain is at least $(1-\delta)\, g_i\left(s_i, \alpha^{\mu}\right)$.
- So, (average deviation gain)$<$(average incentive strength) implies $g_i\left(s_i, \alpha^{\mu}\right) < (1-\delta)^{-1/2}$.

Proof Idea

Based on **variance decomposition**: to provide incentives, continuation payoffs must vary with signals, and due to discounting this variation must be delivered relatively quickkly.

Variance decomposition gives a recursive bound, even though equilibria are not recursive.

3 steps:

1. If not profitable to manipulate in period t, conditional variance of period $t+1$ continuation payoff must be high compared to ratio of period-t deviation gain and detectability (IC+Cauchy-Schwarz).
2. Use law of total variance to apply this lower bound recursively, show discounted sum of payoff variances exceeds a discounted sum of bounds ($(1-\delta)\,\times$ratio of deviation gain and detectability).
3. Use Jensen to convert discounted sum of inequalities to an inequality for the occupation measure, take square root.

Repeated Games with Private Monitoring

In many repeated games, signals are privately observed.

- Stigler 1964: firms choose prices and observe own sales, which depend on all prices + random demand shock.
- Levin 2003, Fuchs 2007: repeated principal-agent relationships where the agent's production is assessed subjectively by the principal.
- "Community enforcement": players interact with different partners each period, only observe outcomes of own relationships.
  - Special observation structure leads to distinct analysis.

Even with public monitoring, the recursive structure exploited by APS+FLM depends on restricting attention to PPE.

- If consider **all** sequential equilibria with public monitoring—where players can accumulate relevant private information as a result of mixing—the situation is similar to that with private monitoring.

Private Monitoring: Preview

Key difficulty with private monitoring: there is no large class of equilibria with a recursive structure, like PPE.

Given this, the literature takes different approaches:

1. Analyze equilibria without a recursive structure, keeping track of different players' beliefs. ("Belief-based approach.")
2. Focus on the (fairly small) class of equilibria where players' beliefs don't matter, which recovers a recursive structure. ("Belief-free approach.")
3. Consider games with private signals but public communication, which allows an analysis similar to PPE (Compte 1996, Kandori Matsushima 1998),
4. Consider games with "almost public" monitoring, which also allows an analysis similar to PPE (Mailath Morris 2002, 2008),

We give a quick overview of these approaches.

Example

Many ideas can be given in the context of a simple 2-period example (MS Ch. 12):

Period 1 stage game is the partnership game/prisoner's dilemma (PD):

	C	D
C	$2, 2$	$-1, 3$
D	$3, -1$	$0, 0$

Period 2 stage game is a coordination game:

	G	B
G	$3, 3$	$0, 0$
B	$0, 0$	$1, 1$

Point: period 2 is short-hand for the continuation game, where the players can coordinate on good or bad continuations (or might mis-coordinate).

Perfect/Public/Almost-Public Monitoring

With perfect monitoring, "grim trigger" (C, then G if CC) is a SPE.

With a public, almost-perfect signal, it's a PPE.

By continuity, with an almost-public (highly correlated), almost-perfect signal, it's a sequential eqm for each player to take C, then G if own signal is CC.

- Since signals are highly correlated, when own signal is CC, w/ high prob opponent's signal is also CC, so optimal to take G.
- Mailath-Morris 02, 06 study repeated games with almost-public monitoring, showing (roughly) that **strict** PPE with **bounded recall** are robust to almost-public monitoring.
- Strictness implies remains best response after slightly perturbing beliefs. Bounded recall implies slightly perturbing monitoring perturbs beliefs only slightly.

Conditionally Independent Monitoring

Almost-public monitoring is important for understanding robustness of public monitoring results to introducing a small chance players get different signals. But it's very special.

Consider instead the opposite extreme of **conditionally independent monitoring**: $p\left(\left(y_1, \ldots, y_N\right) | a\right) = \prod_{i=1}^{N} p_i\left(y_i | a\right)$.

- E.g., $p_i\left(y_i = \bar{y} | a\right) = \begin{cases} p & \text{if } a = CC \\ q & \text{otherwise} \end{cases}$ independent across i

Conditionally Independent Monitoring (cntd.)

Theorem

With conditionally independent monitoring, for any $0 \leq q \leq p < 1$*, in any strict NE, the players take DD in period 1.*

- If $q = 0$ and $p = 1$, grim trigger is clearly a strict NE. So, a discontinuity at $p = 1$.
- If $q = 0$ and $p \approx 1$, grim trigger is a strict NE if y_1 and y_2 are highly correlated. So, a stark difference between almost-public and conditionally independent.

Proof

Let a_1 denote the pure period 1 action profile.

By conditional independence, $\Pr\left(y_{-i} = \bar{y} | a, y_i\right) = \Pr\left(y_{-i} = \bar{y} | a\right)$ for all y_i.

So $\Pr\left(a_{-i,2} = G | a, y_i\right)$ is independent of y_i.

So, by strictness, $a_{i,2}$ is independent of y_i.

- Intuitively, if i sees $\underline{y}$, attributes this to observation noise and ignores it.
- Observation noise can be very unlikely, but a deviation by $-i$ is prob 0!

But then $-i$ takes a myopic best reply in period 1.

A Generalization (Matsushima 91)

Say that σ_i satisfies **independence of irrelevant information (III)** if, whenever two period t histories $h_i^t, \tilde{h}_i^t$ satisfy

$$\Pr\left(h_{-i}^t | h_i^t\right) = \Pr\left(h_{-i}^t | \tilde{h}_i^t\right) \text{ for all } h_{-i}^t,$$

the continuation strategies following h_i^t and $\tilde{h}_i^t$ are identical: $\sigma_i|_{h_i^t} = \sigma_i|_{\tilde{h}_i^t}$.

Theorem

Assume that the stage game has a unique NE, which is in pure strategies, and the monitoring structure satisfies conditional independence and full support. Then the unique repeated game NE such that strategies are pure and satisfy III is the infinite repetition of the static NE.

- Proof is the same as in the 2-period example: argue by induction starting from period 1, and note that where we appealed to strictness, really just used III.

Escape Routes

There are two escape routes from Matsushima's impossibility result.

1. **Relax pure strategies** (and/or conditional independence).
   - If player 1 mixes between C and D in period 1 and is more likely to play G following C and B following D, then observing $\underline{y}$ is informative for player 2, so player 2 may respond with G following $\bar{y}$ and B following $\underline{y}$.
   - Typically, each player must mix to justify the other's mixing.
   - Such equilibria can be constructed, and under some conditions they can support cooperation in the repeated PD. However, the constructions are delicate and difficult to generalize. This route is the **belief-based approach** to private monitoring.
   - If signals are correlated then they are informative even under pure strategies. Equilibria that rely on this correlation also fall under the belief-based approach.
   - Key papers include Sekiguchi 1997, Bhaskar Obara 2002.

Escape Routes (cntd.)

2. Relax III.

- Consider repeated PD with conditionally independent monitoring and suppose each player is more likely to take C today after a good signal of the opponent's last-period action, with probs s.t. opponent is alway indifferent between C and D. (Such probs exist in the PD.)
- This violates III, e.g. because in period 2 a player behaves differently after good and bad period 1 signals, even though these signals carry no information about the opponent's period 1 behavior (C) or signal (indep of own signal).
- This is an example of a **belief-free** equilibrium: both C and D are optimal regardless of the opponent's history, so a player's belief about the opponent's history is irrelevant.
- The belief-free approach is usually more tractable than the belief-based approach. It has been applied/extended to prove the folk theorem for very general private monitoring structures.

Escape Routes (cntd.)

- Key papers on cooperation/folk theorem with belief-free and related methods: Piccione 2002, Ely Valimaki 2002, Matsushima 2004, Ely Horner Olszewski 2006, Horner Olszewski 2006, Sugaya 2022.
- However, the robustness/realism of belief-free strategies is debatable, because they violate III.
- For example, suppose players get small iid taste shocks each period. Then shouldn't the taste shocks determine which action they take when otherwise indifferent, rather than irrelevant past information?
  - Formally, belief-free equilibria may not be **purifiable**. This is a subtle issue, not fully understood. See, e.g., Bhaskar Mailath Morris 2008.

Belief-Free Equilibria in the PD

Consider the PD:

	C	D
C	$1,1$	$-l, 1+g$
D	$1+g, -l$	$0,0$

Assume "ε-perfect conditionally independent monitoring":
$y_i \in \{\underline{y}, \bar{y}$ with

$$p_i\left(\bar{y}|a\right) = \begin{cases} 1-\varepsilon & \text{if } a_{-i} = C \\ \varepsilon & \text{if } a_{-i} = D \end{cases}.$$

We'll see how to use belief-free equilibria (BFE) to support CC.

Plan: Start by considering **perfect** monitoring.

- We'll construct an eqm where (i) players are always indifferent between C and D, and (ii) C is always played on-path.
- When perturb to almost-perfect monitoring, players remain indifferent between C and D, and C is almost always played on path.

Perfect Monitoring

Let each player have 2 states: Good and Bad.

- In Good, play C. If see $\bar{y}$, stay in Good. If see $\underline{y}$, go to Bad with prob p (to be determined).
- In Bad, play D. If see $\underline{y}$, stay in Bad. If see $\bar{y}$, go to Good with prob q (to be determined).

Let V^G, V^B be a player's value when the *opponent's* state is Good,Bad. With perfect monitoring, for players to be indifferent between C, D in both states, we must have

$$\begin{aligned} V^G &= (1-\delta)(1) + \delta V^G = (1-\delta)(1+g) + \delta \left(pV^B + (1-p) V^G \right) \\ V^B &= (1-\delta)(0) + \delta V^B = (1-\delta)(-l) + \delta \left(qV^G + (1-q) V^B \right). \end{aligned}$$

Perfect Monitoring (cntd.)

Solving for V^G, V^B, p, q gives

$$V^G = 1, \quad V^B = 0, \quad p = \frac{1-\delta}{\delta} g, \quad r = \frac{1-\delta}{\delta} l.$$

- This is a valid solution if $\delta \geq \max\left\{\frac{g}{1+g}, \frac{l}{1+l}\right\}$.
- In this case, for any pair $(v_1, v_2) \in [0,1] \times [0,1]$, we can obtain a SPE with payoffs expected payoffs (v_1, v_2) by specifying that at the beginning of the game player 2 randomizes between starting in Good (w/ prob v_1) and Bad (w/ prob $1 - v_1$), and player 1 randomizes between starting in Good (w/ prob v_2) and Bad (w/ prob $1 - v_2$).

Amost-Perfect Monitoring

Now consider the same system of equations for V^G, V^B, p, q, but with almost-perfect monitoring.

The system has a unique solution with perfect monitoring and is linear in the probabilities, so it also has a unique solution for any nearby imperfect monitoring structure.

Moreover, as monitoring imperfections vanish, the solution converges to that with perfect monitoring, so $V^G \to 1$ and $V^B \to 0$.

This implies that, for any pair $(v_1, v_2) \in (0, 1) \times (0, 1)$, there exists a sequential eqm yielding these payoffs under any ε-perfect monitoring structure.

- Formally, ε-perfect monitoring means that for each i there is a partition of Y_i into $\{Y_i(a)\}_{a \in A}$ s.t., for all $a \in A$,

$$\sum_{y_i \in Y_i(a)} p_i(y_i | a) > 1 - \varepsilon.$$

"Standard Repeated Games" vs. "Community Enforcement"

Standard repeated games assume a fixed finite set of players, a well-defined start date, a common notion of calendar time since the start date, and monitoring that satisfies full support and identifiability conditions.

- Good model for interactions with small number of sophisticated players and well-defined start date.
- Small cartel that forms+enforces collusive agreement. Long-term partnership or principal-agent relationship.

"Community Enforcement"

In contrast, conventions governing risk-sharing, exchange, resource management, public goods provision often cover larger groups, individuals may enter and exit over time, may not be a well-defined start date, monitoring may be "decentralized" or "network-like."

The branch of repeated games geared toward this type of setting is called **community enforcement**.

- Formally, special classes of repeated games with imperfect (often private) monitoring.
- New theory issues raised by "decentralized" monitoring or interaction, like random matching or network structure.
- Connects to different areas of economics like political economy, organizations, networks.

There are several classes of community enforcement models.

- Today: repeated games with uniform random matching.
- See survey article "Cooperation in Large Societies" on my webpage for a survey.

Repeated Games with Random Matching

- Each period, population of N players (even) breaks into pairs uniformly at random to play a 2-player stage game.
- Each player perfectly observes their partner's action at the end of each period, but learns nothing about actions taken in other matches.
  - Monitoring where each action is observed either perfectly or not at all are called "partial," "semi-standard," or "network."
  - Intuitively, "opposite" of imperfect public monitoring.

The model comes in two flavors:

- **Anonymous:** do not observe partner's identity before taking action.
- **Non-anonymous:** do observe partner's identity before taking action.

Repeated PD with Anonymous Random Matching

A canonical model of community enforcement is the repeated PD with anonymous random matching.

- Studied in important papers by Kandori 1992, Ellison 1994.

Ultra-low information benchmark model: in reality people usually know something about who they're playing with and what that person's history looks like.

- What's interesting is that cooperation can be possible *despite* ultra-low information.
- Intuition: can't identify deviator, but can provide incentives via *collective punishment*: punish everyone following any defection.

Stage game: with $g, l > 0$,

	C	D
C	$1, 1$	$-l, 1+g$
D	$1+g, -l$	$0, 0$

Cooperation in Nash Eqm (Kandori 1992)

Theorem

In the repeated PD with anonymous random matching, there exists $\bar{\delta} < 1$ such that, for every $\delta > \bar{\delta}$, there is a NE where all players take C in every period along the eqm path.

- Consider **contagion strategies** (=grim trigger): take C until you see anyone take D, then switch to D forever.
- For fixed N, once contagion starts, it spreads throughout the population in finite time with high probability.
- So, for high enough δ, deviating to D at a history when you've only seen C—which starts contagion—is unprofitable.
- These are the only on-path histories, so contagion strategies give a NE.

Remarks

Contagion "punishes everyone" following any deviation, so don't need to identify initial deviator.

In large populations, the order of limits between N and δ matters. If fix δ and take $N \to \infty$, the eqm breaks down because the prob that your action starts affecting your future partners' actions toward you in a relevant timescale goes to 0.

However, since contagion spreads (almost) exponentially, the eqm holds up for quite large populations: if vary δ and N together, contagion strategies give a NE whenever $(1-\delta)\log N \to 0$.

But a problem is that contagion strategies do **not** give a **sequential** eqm: a patient player who observes D in period 1 may want to deviate to C for several periods to slow down the spread of contagion. The punishment for taking D is too harsh, so people won't take it even when prescribed. However, this can be overcome by moderating the punishment...

Cooperation in Sequential Eqm (Ellison 1994)

Theorem

In the repeated PD with anonymous random matching, there exists $\bar{\delta} < 1$ such that, for every $\delta > \bar{\delta}$, there is a **sequential eqm** *where all players take C in every period along the eqm path.*

- Suppose we lessen the severity of contagion so that a player becomes exactly indifferent between C and D in period 1.
- Can do this by using public randomization (if available) to periodically restart cooperation, or by using **threading/Ellison trick** to reduce effective δ.
- **Key observation:** if a player is indifferent between C and D when everyone else is taking C, strictly prefers D once someone else is taking D.
  - Benefit of taking C is it prevents others from switching to D. If others have already switched, benefit is lower.
- So, "forgiving contagion strategies" are a SE.

Theorem shows that anonymity alone doesn't preclude cooperation. But we'll see some qualifications shortly...

Contagion Beyond the PD

Contagion strategies do **not** immediately generalize beyond the PD: players won't switch to a non-dominant "punishing action" if they believe that the rest of the population has not yet switched.

- Nash reversion via contagion strategies only works when the Nash reversion action is dominant in the stage game.

Deb González-Díaz 2019 extend contagion strategies to a larger class of games by having players switch to the punishment phase only after some delay, which is coordinated based on calendar time.

- Significantly more complicated construction, but shows that at least some of the spirit of contagion strategies extends beyond the PD.

Individualized Incentives Under Anonymity

Contagion strategies are simple, but they're not the only way to provide incentives, even under anonymity.

- Suppose player 1 is supposed to take D for T consecutive periods, while everyone else is supposed to take C.
- If a second player deviates to taking D during this T period block, the deviation is statistically detectable by the other players, and so can be punished.

Deb Sugaya Wolitzky 2020 use such asymmetric behavior, together with "block belief-free" ideas following Hörner Olszewski 2006 and Sugaya 2022, to prove the full folk theorem for repeated games with anonymous random matching.

- Result allows both general stage games (beyond PD) and asymmetric target payoffs (e.g., player 1 gets to take D while others take C).
- However, strategies are more complex and less interpretable than contagion strategies. Also, unlike contagion strategies the proof requires extremely high δ.

Robustness?

At first glance contagion strategies seem very fragile: one mistaken D eventually takes over the whole population.

However, Ellison's forgiving contagion strategies have the extra benefit of being more robust. Considering introducing "ε noise": an ε chance that a player is constrained to take D, iid across players and periods. Ellison proves:

Theorem

There exists $\bar{\delta}$ such that, for every $\delta > \bar{\delta}$, there is a strategy profile $s^*(\delta)$ s.t.

1. *There exists $\bar{\varepsilon} > 0$ such that, for every $\varepsilon < \bar{\varepsilon}$, the strategy "'follow $s^*(\delta)$ unless hit by the ε noise" (denoted $s^*(\delta, \varepsilon)$) is a sequential eqm of the game with ε noise.*
2. $\lim_{\varepsilon \to 0} \lim_{\delta \to 1} u(s^*(\delta, \varepsilon)) = 1$.

Proof

- Use a slightly harsher punishment than in the $\varepsilon = 0$ case, so for small $\varepsilon > 0$ C remains optimal when everyone else is taking C, D remains optimal when someone else is taking D.
- A subtlety: after 1000 periods of seeing C with no reset, do you start worrying that contagion has started but you just haven't noticed it yet?
  Turns out not: time passing is bad news, but seeing C's is good news, and your belief $\Pr(\text{contagion has started})$ in the limits after observing many C's turns out to be continuous in ε.
- Efficiency in the iterated limit $\delta \to 1$ then $\varepsilon \to 0$ is easy: $\mathbb{E}[\text{length of punishment phase}]$ that maintains indifference is bounded independent of δ, so as $\varepsilon \to 0$ the prob that contagion gets triggered before each reset goes to 0.

Remarks on Robustness

However, this robustness result has some limitations.

1. Large populations: Community enforcement is motivated by large populations. We saw that contagion strategies remain a NE (and forgiving contagion strategies remain a SE) if $(1-\delta)\log N \to 0$, which is good. But the robustness result only applies if we fix N and then take $\varepsilon \to 0$. It doesn't apply if we fix the expected number of errors in the population $N\varepsilon$, much less if we fix ε and take $N \to \infty$.

2. Commitment types: Robustness result concerns iid noise. Also realistic to think that large populations may contain some "bad types" who always take D. This turns out to be a problem not just for contagion strategies, but for any strategies under anonymity, and even to some extent under non-anonymity.

- Studied in Sugaya Wolitzky 2020, 2021.

Robust Community Enforcement in Large Populations

For these reasons, cooperation with completely anonymous players is not really realistic in a very large population.

In reality, we often do cooperate with rematching in large-population settings, but crucially we are not completely anonymous.

- E.g., cooperation in online platform markets like Uber or AirBnB.

Thus, another important strand of the community enforcement literature studies cooperation in a continuum population where players carry some kind of **record** of their past behavior.

- E.g., Okuno-Fujiwara Postlewaite 1995, Takahashi 2010, Heller Mohlin 2018, Clark Fudenberg Wolitzky 2021.