14.123 Microeconomic Theory III
Spring 2009

MIT 14.123 (2009) by Peter Eso
Lecture 12: Repeated Games

1. Finitely Repeated Games
2. Perfect Folk Theorem
3. Renegotiation Proofness

Read: FT 5.1, 5.2, 5.4; Farrell & Maskin (GEB 1989)

1. Repeated Prisoners' Dilemma

	C	*D*
C	1,1	-1,2
D	2,-1	0,0

- Unique Nash equilibrium $(D,D) \rightarrow (0,0)$. Pareto-optimal (C,C) is <u>not</u> an equilibrium.
- Finite repetition, $t = 1, \ldots, T$: The only <u>Nash</u> outcome is (D,D) in every period.
  - By induction (similar, not ↔ to backward induction, SPE). In any Nash equilibrium σ^*, both players play D in period T. Hence for any history that has positive probability up to T-1, player i has no incentive to play C at T-1, because no matter what he does his opponent plays D in period T anyway. Induction on the number of periods gives the result. ■
- In experiments (with humans or in Axelrod's tournament) we see cooperation. "Tit-for-tat" does well in reality with $T < \infty$.

Single-Deviation Principle

- Repeated games belong to the class of multi-stage games with observable actions ("almost-perfect information games").
- THM: A strategy profile of a multi-stage game with observable actions (finite-horizon or infinite-horizon with continuity at ∞) is a subgame-perfect equilibrium (SPE) iff the following holds:

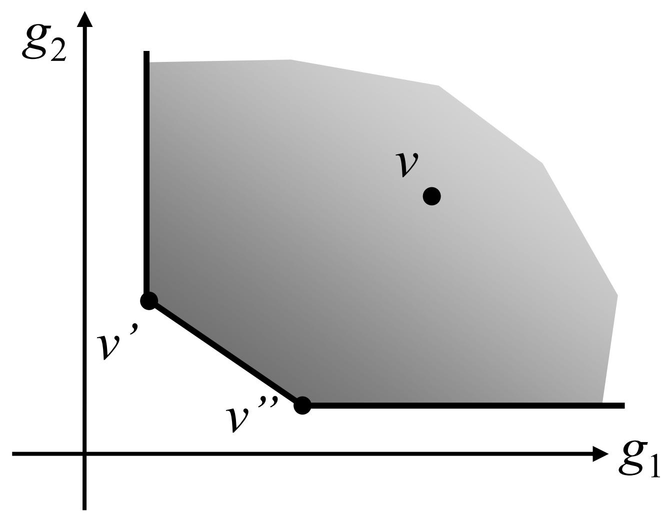
> For any history h^t (=the play up to, not including t) and i assume
>
> – at t and thereafter everybody except for i plays according to the proposed equilibrium strategy profile, and
>
> – at t+1 and thereafter i plays the proposed strategy profile;
>
> then i does not have an incentive to deviate at h^t.

SPE with Finite Repetition

- Set of SPE may expand even with finite repetition (not in PD).
- Ingredients: Multiple equilibria that the players rank differently, sufficiently long time horizon, and patience.
- THM (Benoit and Krishna, 1985); two players, no discounting.

Suppose (v_1', v_2') and (v_1'', v_2'') are stage-game Nash eqm payoffs with $v_1' > v_1''$ and $v_2'' > v_2'$.

$\forall (v_1, v_2)$ feasible & in the shaded area, $\forall \varepsilon > 0$, there is $\underline{T} < \infty$ such that G^T with $T \geq \underline{T}$ has SPE with average payoffs within ε of (v_1, v_2).



Proof

- Choose t^* such that

$$t^*(v_1'' - v_1')/2 > w_1 \equiv \max_a g_1(a),$$

$$t^*(v_2' - v_2'')/2 > w_2 \equiv \max_a g_2(a).$$

- Proposed SPE, at least $2t^*$ periods before the end of the game:

  A. Play (v_1,v_2) until time $T - t^*$ unless someone deviates.

  B. If no deviation in (A), then in the final t^* periods alternate between (v_1',v_2') and (v_1'',v_2'').

  C. If P1 deviates in (A), then play (v_1',v_2') to the end.
  If P2 deviates in (A), then play (v_1'',v_2'') to the end.

- Indeed approximates (v_1,v_2) for T sufficiently large.

Proof, continued

- Why SPE?
- Denote $t > t^*$ the remaining time.
- If no-one deviated before, P1 gets payoff $(t - t^*)v_1 + t^*(v_1' + v_1'')/2$ if conforms, at most $w_1 + (t\text{-}1)v_1' < w_1 + tv_1'$ if deviates. Difference: $(t - t^*)(v_1 - v_1') + t^*(v_1'' - v_1')/2 - w_1 > 0$ for conform.
- Same goes for P2 if no-one deviated before.
- If anyone deviated already, then Nash equilibrium is played in every period, subgame perfect.
- In the final t^* periods, alternate over two Nash equilibria: SPE. ■

2. Infinite Repetition

- Repetition without known bound (can be finite in expectation) expands the set of equilibria even in the Prisoners' Dilemma.
- THM: Infinitely repeated PD, discounted payoffs with $\delta > ½$: "Grim Trigger" (=play C as long as both play C, play D forever if any player ever plays D) is SPE and yields (C,C), $\forall t$.
  - ■ Equilibrium payoff is 1 per period. Single-period deviation yields payoff 2, and 0 from then on. $1/(1-\delta) > 2$ for $\delta > ½$. ■
- This construction is rather special: In the Prisoners' Dilemma players can punish with stage game Nash equilibrium. This makes the infinitely repeated game equilibrium SPE.
- If the punishment is itself not an equilibrium (=not credible), then the repeated-game equilibrium is only Nash, not SPE.

Destruction By Repetition

- In the one-shot game the unique Nash equilibrium is (A,A) because A is strictly dominant.
- (A,A) in all periods is SPE for finite or infinite repetition.

	A	B	C
A	2,2	2,1	0,0
B	1,2	1,1	-1,0
C	0,0	0,-1	-1,-1

- Claim: Infinite repetition with $\delta > ½$: (B,B) $\forall t$ is SPE outcome.

■ Strategy $s^* =$ {Play B at $t = 1$ and $\forall t$ such that both players played s^* in period t–1; play C at t if someone deviated from s^* at t–1}.

- If $s^*(h_t) = B$: Using s^* get $K + \delta^t + \delta^{t+1} + \ldots$; one-shot deviation to A yields $K + 2\delta^t - \delta^{t+1} + \delta^{t+2} + \ldots$. Gain is $\delta^t (1 - 2\delta) < 0$.
- If $s^*(h_t) = C$: Using s^* get $K - \delta^t + \delta^{t+1} + \ldots$; one-shot deviation to A or B yields $K + 0\delta^t - \delta^{t+1} + \delta^{t+2} + \ldots$. Gain is $\delta^t (1 - 2\delta) < 0$. ■

General Notation

- Each period play stage game g; infinitely repeated game is g^∞. In g, players are $N = \{1,\ldots,n\}$, actions $a_i \in A_i$ for $i = 1,\ldots,n$.
- $g_i(\alpha)$ is i's stage game payoff given a (mixed) action profile α.
- σ_i is infinitely-repeated game strategy for player i. Specifies (mixed) action α_i for all histories $h^t = (a^0,\ldots,a^{t-1})$, $\forall t \geq 0$.
- $v_i(\sigma) = (1-\delta)\sum_{t\geq 0} \delta^t \sigma(h^t)\, g_i(\alpha^t|\sigma,h^t)$ is average discounted payoffs of strategy-profile σ. Comparable to per-period payoff.
- If the period-0 actions are already known, one can rewrite this as $v_i(\sigma) = (1-\delta)g_i(a^0) + \delta v_i(\sigma^c(a^0))$, where $\sigma^c(a^0)$ is the strategy profile in periods $t = 1,2,\ldots$ induced by σ given period-0 actions a^0.
- $S(\sigma)$ = set of continuation profiles of σ after every finite history. Note: σ is SPE of g^∞ iff all $\sigma' \in S(\sigma)$ is SPE of g^∞.

Payoff Constraints In Any NE

- Here are two results regarding on the set of average discounted payoffs that may be the result of a Nash equilibrium of $g^T(\delta)$:
- OBS 1: Feasibility. If $(v_1,\ldots,v_n)$ are the average discounted payoffs in a Nash equilibrium, then

  $(v_1,\ldots,v_n) \in \text{co}\{(x_1,\ldots,x_n) \mid \exists(a_1,\ldots,a_n) \text{ with } x_i = g_i(a_1,\ldots,a_n), \forall i\}.$
- DEF: Minmax payoff, $\underline{v}_i = \min_{\sigma_{-i}} \max_{\sigma_i(\sigma_{-i})} g_i(\sigma_i(\sigma_{-i}), \sigma_{-i})$.
- OBS 2: Individual Rationality. If $(v_1,\ldots,v_n)$ are the average discounted payoffs in a Nash equilibrium, then $v_i \geq \underline{v}_i$ for all i.

■ Suppose $(\sigma^*_i,\sigma^*_{-i})$ is NE of g^T, and construct σ_i so that $\sigma_i(h^t)$ is a best-response to $\sigma^*_{-i}(h^t)$ at every history h^t. Then,

$$U_i(\sigma^*_i,\sigma^*_{-i}) \geq U_i(\sigma_i,\sigma^*_{-i}) \geq (1-\delta)/(1-\delta^{T+1}) \left(\sum_t \delta^t\, \underline{v}_i\right) = \underline{v}_i\,. \ ■$$

Nash Folk Theorem For g^∞

- <u>THM</u>: If $(v_1,\ldots,v_n)$ is feasible & strictly individually rational, then there exists $\underline{\delta} < 1$ such that $\forall \delta \geq \underline{\delta}$, there is a NE of $g^\infty(\delta)$ with average payoffs $(v_1,\ldots,v_n)$.

- Assume for simplicity, $\exists (a_1,\ldots,a_n) \in A$ with $g_i(a_1,\ldots,a_n) = v_i$.

- Denote m^i_{-i} the strategy-profile of players other than i that hold player i to his minmax payoff and m^i_i a best response to m^i_{-i} .

- Proposed equilibrium strategies: Each i plays
  - a_i at h_t such that $(a_1,\ldots,a_n)$ has been played $\forall t' < t$.
  - m^j_i if player j was the first player to have deviated (or, if multiple players deviated first, simultaneously, then the lowest numbered one among them).

Proof, continued

- If player i follows this strategy, then his average payoff is v_i.
- If player i deviates in period t, then his average payoff is at most

$$(1\text{-}\delta)(v_i + \ldots + \delta^{t-1} v_i + \delta^t w_i + \delta^{t+1} \underline{v}_i + \delta^{t+2} \underline{v}_i + \ldots),$$

where $w_i = \max_{a \in A} g_i(a)$ is i's highest feasible payoff in G.

- Deviation is not worth it if

$$(w_i - v_i) \leq \delta/(1\text{-}\delta)\ (v_i - \underline{v}_i).$$

- Choose $\underline{\delta}$ such that $\underline{\delta}/(1\text{-}\underline{\delta}) \geq \max_i\ (w_i - v_i) / (v_i - \underline{v}_i)$. ■

- The theorem is useful as it characterizes the set of all Nash equilibria of $g^\infty(\delta)$, at least for high enough δ.

Why Go Beyond Nash

- Nash equilibrium is not a particularly appropriate concept for dynamic games. Reason: Incredible punishment threats.

- We can sustain (*C*,*C*) in the infinitely-repeated game by P2 punishing P1 forever in case P1 ever deviates to *D*.

	C	*D*
C	1, 1	0, -10
D	2, 1	0, -10

- But the punishment hurts P2 more than it hurts P1; P2 may not want to carry it out.

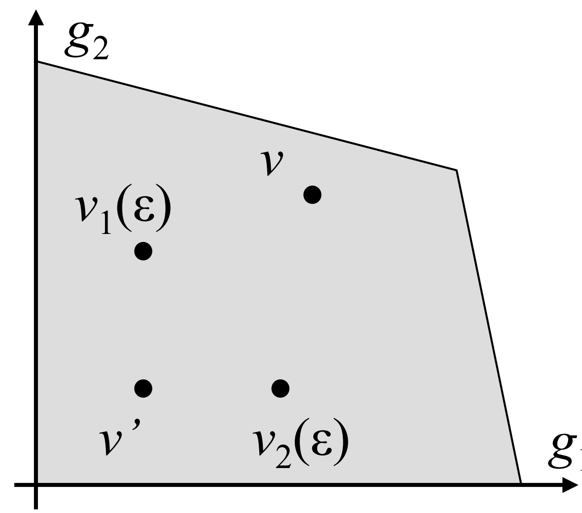
- The example calls for requiring <u>subgame perfection</u>.

Perfect Folk Theorem

- <u>THM</u> Fudenberg and Maskin (1986). Let V^* be the set of feasible and strictly IR payoffs of G. Assume $\dim(V^*) = n$. Then, for any $(v_1,\ldots,v_n) \in V^*$ there exists $\underline{\delta} < 1$ such that for all $\delta \geq \underline{\delta}$, there is a SPE of $g^\infty(\delta)$ with average payoffs $(v_1,\ldots,v_n)$.

■ Wlog denote $\underline{v}_i = 0$, moreover assume $\exists a \in A$: $g_i(a) = \underline{v}_i$ for all i.

- Pick $v' \in \mathrm{int}(V^*)$ with $v'_i < v_i$ for all i.
  Let T such that $Tv'_i > w_i = \max_{a \in A} g_i(a)$.
- Pick $\varepsilon > 0$ so that for each i,
  $v_i(\varepsilon) = (v_i', v'_{-i}+\varepsilon) \in V^*$ and $v_i'+\varepsilon \leq v_i$.
  Let a^i such that $g_i(a^i) = v_i(\varepsilon)$.



Proof, continued

- Denote m^i the strategy-profile that minmaxes player i. Assume that m^i is either pure, or mixing probs can be detected.
- Here is the proposed SPE. Each player i plays the following strategy, which prescribes behavior for three "phases".

I. Play $(a_1,\ldots,a_n)$ as long as no-one deviates from $(a_1,\ldots,a_n)$.
If player j deviates from phase I then go to phase II_j.

II_j. Play m^j_i for T periods, then go to phase III_j if no-one deviates.
If player k deviates in II_j, then start over II_k.

III_j. Play a^j_i as long as no-one deviates from III_j.
If player k deviates in III_j, then go to phase II_k.

Proof, finished

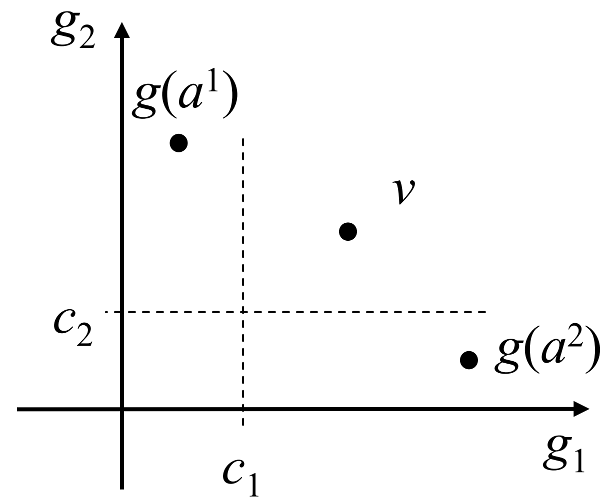
- Single-deviation principle in each phase.
- In phase I, deviating once yields at most $(1\text{-}\delta)w_i + \delta^{T+1}v'_i$ which is less than $v_i = (1\text{-}\delta^{T+1})v_i + \delta^{T+1}v_i$ if δ is close to 1, e.g., $\delta > (1+1/T)^{1/T}$, as $(1\text{-}\delta^{T+1})v_i = (1\text{-}\delta)(1+\delta+\ldots+\delta^{T})v_i > (1\text{-}\delta)Tv_i > (1\text{-}\delta)w_i$.
- In phase II_i, deviation by i postpones everything by T, not worth it.
- In II_j, if i deviates, he gets $(1\text{-}\delta)w_i + \delta^{T+1}v'_i$; if he conforms when K periods are still left of II_j, he gets $(1\text{-}\delta^{T+1\text{-}K})g_i(m^j) + \delta^{T+1\text{-}K}(v'_i+\varepsilon)$. Conform iff $\delta^{T+1}\varepsilon \geq (1\text{-}\delta)w_i + (1\text{-}\delta^{T+1\text{-}K})g_i(m^j) + (\delta^{T+1\text{-}K}\text{-}\delta^{T+1})(v'_i+\varepsilon)$, which holds as δ approaches 1 (LHS $\rightarrow \varepsilon$, RHS $\rightarrow 0$).
- In phase III_i or III_j the proof is like in phase I: Deviation provides gains for one period, loss for T periods, not worth it. ■

3. Renegotiation Proofness

- Criticism of repeated-game SPE with "punishment phases": Players may want to renegotiate, if both are hurt by punishment. Farrell & Maskin GEB'89 propose to consider the following.
- DEF: An SPE of g^{∞}, σ, is Weakly Renegotiation Proof (WRP), if $\forall \sigma', \sigma'' \in S(\sigma)$, σ' does not strictly Pareto-dominate σ''.
- Think of $S(\sigma)$, all possible infinite strategy profiles induced by σ, as "the plays we agree are in the playbook". If $\sigma' \in S(\sigma)$ strictly Pareto-dominates $\sigma'' \in S(\sigma)$, then the players renegotiate σ' to σ''.
- In PD, "(D,D) forever" has unique continuation, hence it is WRP. "Grim Trigger" is not WRP; it dominates continuation after (D,D).
- Internal consistency, not comparison across SPE's.

Theorem (Farrell & Maskin '89)

- Consider two players; normalize minmax payoffs to 0 and let V^* denote all feasible, IR payoffs.
- Suppose $(v_1,v_2) \in V^*$. If there exist actions $(a^1{}_1,a^1{}_2)$, $(a^2{}_1,a^2{}_2)$ such that



(1) $c_1 \equiv \max_x g_1(x,a^1{}_2) < v_1$, $g_2(a^1) > v_2$

(2) $c_2 \equiv \max_x g_2(a^2{}_1,x) < v_2$, $g_1(a^2) > v_1$

then for δ near 1 there is a WRP equilibrium with payoffs (v_1,v_2).

* Conversely, if σ is WRP equilibrium with payoffs (v_1,v_2), then there exist action-pairs a^1 and a^2 satisfying (1) & (2) <u>weakly</u>.

Proof

- First, we construct a WRP equilibrium if (1) and (2) hold.
- Suppose $(v_1, v_2) = g(a_1, a_2)$. Propose WRP equilibrium as follows:

(I): Play (a_1, a_2) until i deviates; then go to (II_i).

(II_i): Play a^i for t_i periods, such that $t_i\, g_i(a^i) + w_i < (t_i + 1)\, v_i$.
Then go back to (I). If j deviates from II_i then (re)start II_j.

- t_i exists by $g_i(a^i) < v_i$ and makes deviation from (I) unprofitable.
- Set δ high enough so that $p_i = (1 - \delta^{t_i})\, g_i(a^i) + \delta^{t_i}\, v_i$ satisfies

$$p_i > c_i \quad \text{and} \quad (1 - \delta)w_i + \delta p_i < v_i.$$

Possible because $g_i(a^i) \leq c_i < v_i$.

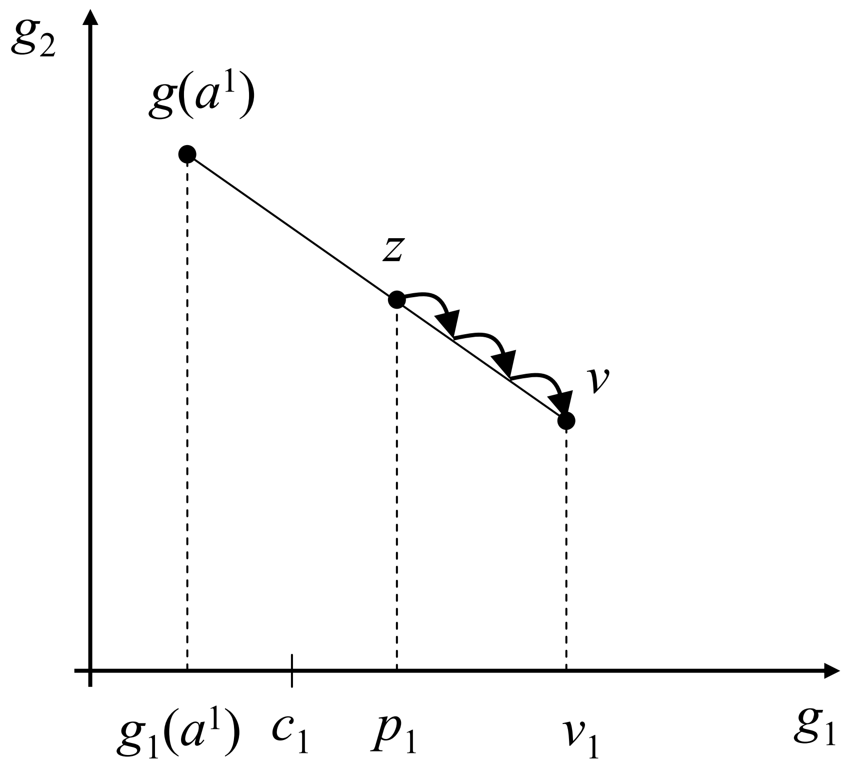
- Claim: Proposed strategies form WRP eqm for such high δ.

Illustration for i = 1

- If P1 deviates from (I), then (II_1) prescribes t_1 periods of $g(a_1)$, and then v forever; payoffs are

  $z = (1 - \delta^{t_1})g(a^1) + \delta^{t_1} v_1$.

- During (II_1), slide to v.
- If P1 deviates in (II_1), he gets $(1 - \delta)c_1 + \delta p_1 < p_1$.
- P2 does not deviate from (II_1) because $g(a^1) > v_2$.



- Continuation payoffs lie between z and v, Pareto-unranked, WRP!

*Proof Still Not Over

- Second: Given δ, if WRP eqm with payoffs (v_1,v_2) exists, then there are actions a^1, a^2 such that (1) and (2) hold weakly.
- We show that a^1 satisfying (1) weakly exists; a^2 & (2) analogous.
- Let σ be the WRP eqm given δ. If there is an action-pair a such that $g_1(a) = v_1$ and $g_2(a) \geq v_2$, and in addition, $\max_x g_1(x,a_2) \leq v_1$ as well, then a itself satisfies (1).
- Otherwise, consider σ^1, the worst continuation of σ for P1 after period 1 (prompted by a first-period action a' with $g_1(a') \geq v_1$). If there are multiple worst-continuations of σ, then take the one that is best for P2.
- a^1 = initial action of σ^1. We claim it satisfies (1) weakly.

*Proof Finished

- The worst continuation of σ for P1, σ^1, satisfies $g_1^*(\sigma^1,\delta) \leq v_1$ and $g_2^*(\sigma^1,\delta) \geq v_2$. (The former by def, the latter by WRP.)
- $g_2(a^1) \geq g_2^*(\sigma^1,\delta)$ ($\geq v_2$), establishing the second inequality in (1), because $g_2(a^1) < g_2^*(\sigma^1,\delta)$ would imply that $\underline{\sigma}^1$, the continuation of σ^1 after a^1, satisfies $g_2^*(\underline{\sigma}^1,\delta) > g_2^*(\sigma^1,\delta)$, hence by WRP $g_1^*(\underline{\sigma}^1,\delta) \leq g_1^*(\sigma^1,\delta)$, contradicting that σ^1 is the worst continuation for P1.
- The first inequality in (1), weakly, is that $\max_x g_1(x,a^1{}_2) \leq v_1$.
  We show a bit more: $\max_x g_1(x,a^1{}_2) \leq g_1^*(\sigma^1,\delta)$ ($\leq v_1$).
  If $\max_x g_1(x,a^1{}_2) > g_1^*(\sigma^1,\delta)$, then P1 could profitably deviate in the first period of playing σ^1, and since his continuation payoff cannot be lower than $g_1^*(\sigma^1,\delta)$, by definition of σ^1, the deviation would be profitable overall. Hence $\max_x g_1(x,a^1{}_2) \leq g_1^*(\sigma^1,\delta)$. ■