Repeated Games

No tangible link between periods, but players can condition current play on their information about past actions.

This can allow new equilibria because players know have the ability to "reward" or "punish" opponents.

But repeating the game doesn’t get rid of any equilibrium outcomes. **Observation:** *(for both finite-horizon and infinite-horizon repeated games)*:

If $\alpha^*$ is a Nash equilibrium of the stage game (a "static equilibrium") then the constant strategy profile $s^*(h_t) = \alpha^* \ \forall t, h_t$ is an SPE. Moreover if there are multiple static NE, an SPE can play different static NE in different periods.
The point of looking at repeated games is to see what happens when the discount factor is not small.

Hard to characterize equilibrium set for fixed discount factor.

Strongest conclusions emerge when the discount factor is very close to 1, and the game is infinitely repeated.

Here we have the "folk theorems," which show that a large set of outcomes can arise in equilibrium.

There are different versions of the folk theorem for different payoff criteria (discounting or time average) different equilibrium concepts (Nash vs. SPE), different sorts of information assumptions, etc. Also some related theorems that are easier to prove but whose conclusions not quite as strong.
Overview of Repeated Games

a) I will focus on discounted payoffs, also literature on time averaging.

b) Finite vs Infinite Horizon

c) Classic case: fixed set of players. Will discuss “community enforcement” later.

d) Start with observed actions, all LR players. Then relax to imperfect observability, LR and SR; then relax to unknown monitoring structure, private information, stochastic games, etc.
Observed actions

**Theorem** (J. Friedman 1971): Let $\alpha^*$ be a static equilibrium, let $a$ be an action profile with $g_i(a) > g_i(\alpha^*)$ for all players $i$. Then there is a $\delta < 1$ such that for all $\delta \in (\delta, 1)$, there is an SPE in which $a$ is played each period on the equilibrium path.

**Proof:** constructive, use reversion to static equilibrium.

Discount factor plays a key role: It measures the “value” of the lag between cheating and punishment, so shorter time periods correspond to larger discount factors, which makes "collusion" harder to sustain.

Discount factor near 1 corresponds to low time preference and/or short periods- equivalent here but not with imperfect monitoring…
Friedman's theorem: many outcomes are consistent with SPE when the players are sufficiently patient.

The set of SPE is typically larger; "virtually anything" can be supported as a SPE when players are patient- i.e. any feasible payoff vector that Pareto dominates the minmax payoffs.

**Theorem:** (Fudenberg-Maskin 1986) Suppose that a "full-dimensionality condition" is satisfied. Let $a$ be an action profile with $g_i(a) > v_i$ for all $i$. Then there is a $\delta < 1$ such that for all $\delta \in (\delta, 1)$, there is an SPE in which $a$ is played each period on the equilibrium path.

Proof: constructive.
Complicated step is constructing strategies that induce player $i$'s opponents to minmax him, at least for some length of time; we need to be able to do this to force $i$’s equilibrium payoff close to his minmax.

The reason this construction takes some work is that a player's payoff when minmaxing another player can be less than the player's own minmax payoff, in which case minmaxing can't be enforced just by the threat of being minmaxed; the strategies need to use "rewards" as well as "punishments."

This is why the “full dimension” condition is needed, need to be able to design rewards for punishers w.o. rewarding the punishee.
Equilibrium strategies don’t need to use rewards for punishing if there is a “pure strategy mutual minmax action” - an action $a^*$ such that $\max_{a_i} g_i(a_i, a^*_{-i}) \leq v_i$. Then specify strategies that respond to any and all deviations by having $T$ periods of play of $a^*$ - so deviating during a “punishment phase” just restarts the punishment.

Then just need to check that (a) punishment period is long enough that no incentive to deviate on the equilibrium path and (b) punishment period is short enough that each player $i$’s value at start of punishment phase exceeds $v_i$ so don’t want to deviate and prolong punishment.
But with more than 2 players there needn’t be a mutual minmax action—so a player isn’t minmaxed while minmaxing another. Now the strategies need to provide incentives for carrying out a “punishment.” This is easiest to do when the minmax strategies are pure—otherwise even with 2 players the rewards need to depend on the outcome of the randomization so as to make players indifferent.

Hörner-Olszewski *Econometrica* [2006] give a different “mostly constructive” proof (and adapt it to prove a folk theorem for games with private, almost-perfect, monitoring.)
Comments on the constructive approach:

- Fine way to prove folk theorem but less good if constructed set is smaller then need some other upper bound on equilibrium set.

- Constructed strategies aren’t unique equil. with given outcome path.

- As in HO, constructions can be adapted to other settings.
Comments on folk theorem:

- Too many equilibria, would like sharper predictions. But equilibrium selection in repeated games is an open question; Fudenberg Maskin AER [1990] and [2009] unpublished have an evolutionary argument for cooperation in repeated games with unrestricted strategy spaces; also a larger literature on evolutionary game theory applied to repeated PD with small sets of strategies, ie. TfT, grim, WSLS, all D, all C.

<table>
<thead>
<tr>
<th></th>
<th>C</th>
<th>D</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td>2, 2</td>
<td>−1, 3</td>
</tr>
<tr>
<td>D</td>
<td>3, −1</td>
<td>0, 0</td>
</tr>
</tbody>
</table>

*TfT: start C then play what opponent played yesterday; not subgame perfect.*

*WSLS (perfect TfT): start C then C if yesterday (C,C) or (D,D)*
• It’s hard to look at data and see what sorts of strategies players use especially in games with perfect monitoring- don’t see play at many info sets. Older literature either had people enter computer codes (Axelrod) or played against experimenter not other subjects (Roth and Murnigan) or no cash payoffs. Newer experiments: Dal Bó AER [2005]- frequency of cooperation in a prisoner’s dilemma increase in continuation probability; Dal Bó and Frechette AER forthcoming look at how coop varies with payoff matrix and cont. probability.

Here “R” is the reward payoff if both cooperate, and “risk dominant” is in a 2x2 game vs. Tit for Tat.
### Table 1: Stage Game Payoffs

<table>
<thead>
<tr>
<th></th>
<th>C</th>
<th>D</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td>R, R</td>
<td>12, 50</td>
</tr>
<tr>
<td>D</td>
<td>50, 12</td>
<td>25, 25</td>
</tr>
</tbody>
</table>

### Table 2: Cooperation as Equilibrium (SGPE) and Risk Dominant (RD) Action

<table>
<thead>
<tr>
<th></th>
<th>R=32</th>
<th>R=40</th>
<th>R=48</th>
</tr>
</thead>
<tbody>
<tr>
<td>δ =1/2</td>
<td>Neither SGPE or RD</td>
<td>SGPE</td>
<td>SGPE and RD</td>
</tr>
<tr>
<td>δ=3/4</td>
<td>SGPE</td>
<td>SGPE and RD</td>
<td>SGPE and RD</td>
</tr>
</tbody>
</table>
Point: when coop is “just barely” SPE it isn’t so common.
They also try to identify the strategies used: assume everyone uses one of fixed set of strategies, add errors, then compute fraction using each strategy by max likelihood. In high-coop treatment they find 56% TfT a 24% “T2”, 12 % Grim and 7% “All C” but T2 seems implausible to me, may be an artifact.

<table>
<thead>
<tr>
<th>δ \ R \ δ \ R</th>
<th>32</th>
<th>40</th>
<th>48</th>
<th>32</th>
<th>40</th>
<th>48</th>
</tr>
</thead>
<tbody>
<tr>
<td>½</td>
<td>34.09</td>
<td>&lt;*</td>
<td>54.00</td>
<td>&lt;</td>
<td>56.52</td>
<td>½</td>
</tr>
<tr>
<td></td>
<td>=</td>
<td>v</td>
<td>^</td>
<td></td>
<td>v</td>
<td>^</td>
</tr>
<tr>
<td>¾</td>
<td>34.09</td>
<td>&lt;</td>
<td>36.84</td>
<td>&lt;*</td>
<td>56.82</td>
<td>¾</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>δ \ R \ δ \ R</th>
<th>32</th>
<th>40</th>
<th>48</th>
<th>32</th>
<th>40</th>
<th>48</th>
</tr>
</thead>
<tbody>
<tr>
<td>1/2</td>
<td>9.81</td>
<td>&lt;**</td>
<td>18.72</td>
<td>&lt;***</td>
<td>38.97</td>
<td>1/2</td>
</tr>
<tr>
<td></td>
<td>^***</td>
<td>^***</td>
<td>^***</td>
<td></td>
<td>^**</td>
<td>^***</td>
</tr>
<tr>
<td>3/4</td>
<td>25.61</td>
<td>&lt;**</td>
<td>61.10</td>
<td>&lt;*</td>
<td>85.07</td>
<td>¾</td>
</tr>
</tbody>
</table>
Ongoing project with Dreber and Rand: experimental play of PD with errors. (imperfect pub obs.)

How far back to people look? (a bit easier to tell here since more info sets hit, but the econometrics is still complicated.)

What if WSLS isn’t an equilibrium because 2 periods of punishment needed to deter deviation? (we still get cooperation).
Repeated Games with Public Monitoring

Stage game: players $i = 1, \ldots, I$ simultaneously choose actions $a_i \in A_i$. 

$\alpha_i$ denotes a mixed action.

$y \in Y$ are public outcomes; $Y$ is assumed finite except where noted.

Action profile $a \in A \equiv \times_i A_i$ induces a probability distribution on $Y$; the probability of $y$ given $a$ is denoted $\pi_y(a)$.

The expected payoff to an action profile is $g_i(a)$. 
Typically assume \( g_i(a) = \sum_y \pi_y(a)r_i(a_i, y) \),

where player i's realized payoff is \( r_i(a_i, y) \) depends on actions of other players only through \( y \).

(Otherwise if the players observe their own payoffs each period, they would have a source of private information about opponents’ actions. This doesn’t matter for the analysis but could make PPE less interesting.)
Repeated Game:

Public information at $t$ is $h^t = \{y^0, \ldots, y^{t-1}\}$

Player i also has private info at $t$: $z^t_i = \{a^0_i, \ldots a^{t-1}_i\}$.

Strategies for i: maps from histories $(h^t, z^t_i)$ to actions.

Payoffs: $u_i = (1 - \delta)\sum_{t=0}^{\infty} \delta^t g_i(a)$,

where $\delta \in (0,1)$ is the discount factor.

(normalized discounted payoffs.)
(infinite horizon)
Examples:

*Observed actions:* $Y = A$, $\pi_y(a) = I(a)$

*Green-Porter oligopoly model:*
  $a_i$ is firm i’ output, $y$ is price, $\pi_y$ depends only on sum of the outputs

*Repeated partnership* (Radner, Radner-Myerson-Maskin)
  $a_i$ is effort, $y$ is output

*Repeated insurance* (E. Green):
  Each period player i learns endowment $\theta_i$. These are i.i.d. and private information. Players then make possibly false "reports" of their endowment, so action space is $A_i = (\Theta_i)^{\Theta_i}$, the space of all maps from $\Theta_i$ to $\Theta_i$. 
**Definition:** (FLM) Strategy $s_i$ is a *public strategy* if it depends only on public information. That is, player $i$ is using a public strategy if $s_i(h^t, z^t_i) = s_i(\bar{h}^t, \bar{z}^t_i)$ whenever $h^t = \bar{h}^t$.

**Lemma:** Any pure strategy NE is equivalent to a NE in public (and pure) strategies

**Idea:** Suppose $i$ uses a pure strategy. Then in period 0 he plays some $\bar{a}_i^0$. In period 1, he plays according to some $\sigma_i(y^0, a_i^0)$. Replace this with $\sigma_i'(y^0) = \sigma_i(y^0, \bar{a}_i^0)$, which is a public strategy; it has the same outcome, and so leaves best responses of other players unchanged. Intuitively, there is no reason for $i$ to condition on his own past play, as it won’t help him predict the future play of his opponents.
The lemma isn’t true for all mixed-strategy NE: there can be mixed NE that are not equivalent to public equilibria.

Reason: if all players are mixing, then $i$’s belief about $j$’s past play, conditional on the public history, can depend on $i$’s own past play. So if $i$ thinks $j$’s current play depends on $j$’s past play (in addition to the public history), then $i$’s current best response can depend on $i$’s past actions. And the same is true for $j$, so we can construct a NE in which all players condition in this way. (See FT exercise 5.10)
We'd like to refine NE to something like SPE but in general, these games have no proper subgames so SPE is vacuous.

However, if all players use public strategies, it’s as if there were proper subgames:

At any date $t$ players don’t know exactly what actions were played in the previous periods, but if everyone use public strategies this missing information doesn’t matter: We can still define continuation payoffs and test for Nash equilibrium as if it were a proper subgame.
**Definition** Strategy profile $\sigma$ of the repeated game is a *perfect public equilibrium* ("PPE") if

(a) each $\sigma_i$ is a public strategy
(b) for each $t$ and $h'$, the strategies yield a "NE" from that date on.

**Claim**: set of PPE has a "recursive structure":
The set of PPE at date 0 is isomorphic to the set of PPE in continuation games corresponding to $(t,h')$.
So the set of payoffs of PPE equals the set of continuation payoffs of PPE from $(t,h')$.

We will focus on PPE, but these games can have other equilibria as well.
Dynamic Programming and Self-generation

This is the multiplayer version of principle of optimality of dynamic programming, “Unimprovable implies optimal”:

Consider decision problem with finite set of states $K$, actions $a$, transition probabilities $p(k_{t+1} = k' \mid k_t = k, a_t = a)$, and objective $V = \sum_{t=0}^{\infty} \delta^t u(a_t, k_t)$, $u$ bounded.

Then if there are numbers $v^*(k)$ such that for all $k$,

$$v^*(k) = \max_a (1 - \delta)u(a, k) + \delta \sum_{k'} p(k' \mid (a, k))v^*(k'),$$

$v^*(k)$ is the value of following the optimal decision rule when starting in state $k$.

Here the states are endogenous, depend on the equilibrium strategies.
Definition \((\alpha, v)\) is enforceable w.r.t. \(\delta\) and \(W \subseteq R^I\) if there exists a function \(w: Y \rightarrow W\) s.t for all players \(i\)

(i) \(v_i = (1 - \delta)g_i(\alpha) + \delta \sum \pi_y(\alpha)w_i(y)\)

(if play today is \(\alpha\), and continuation payoffs are given by \(w\), then overall expected payoff is \(v\))

(ii) \(\alpha_i\) solves

\[\max_{\alpha_i}(1 - \delta)g_i(\alpha_i, \alpha_{-i}) + \delta \sum \pi_y(\alpha_i, \alpha_{-i})w_i(y)\]

Motivation: in any PPE, actions played in each period are enforced by the equilibrium continuation payoffs
$B(\delta, W, \alpha)$ is the set of all payoffs $v$ such that $(\alpha, v)$ is enforceable w.r.t $(\delta, W)$.

If there exists $\alpha$ such that $(\alpha, v)$ is enforceable w.r.t. $(\delta, W)$, we say that $v$ is generated by $(\delta, W)$.

\[ B(\delta, W) = \bigcup_{\alpha} B(\delta, W, \alpha) \] is the set of all payoffs generated by $(\delta, W)$.

$E(\delta)$ is the set of all PPE payoff vectors for a given $\delta$. 
**Theorem (APS):** $E(\delta) = B(\delta, E(\delta))$.

**Proof:** We need to show $E(\delta) \subseteq B(\delta, E(\delta))$ and $E(\delta) \supseteq B(\delta, E(\delta))$.

If $v \in E(\delta)$ then no player wants to deviate in the first period. And the continuation payoffs lie in $E(\delta)$, so $E(\delta) \subseteq B(\delta, E(\delta))$.

Conversely, fix $v \in B(\delta, E(\delta))$, we must show it is in $E(\delta)$. To do this we will build an equilibrium with payoffs $v$.

Pick an $\alpha$ such that $(\alpha, v)$ is enforceable w.r.t $(\delta, E(\delta))$, and let $w$ be the map with values in $E(\delta)$ that enforces it.

For each $y$ in $Y$, let $\hat{\sigma}(y)$ be a PPE with payoff vector $w(y)$. 

*(this exists because $w(y) \in E(\delta)$.)*
Specify strategies:
Play $\alpha$ at start. If $y$ occurs in stage 0, then play $\hat{\sigma}(y)$ from date 1 on.

That is, $\sigma_i(0) = \alpha_i$
$\sigma_i(y^0) = \hat{\sigma}_i(y^0)(0)$
$\sigma_i(y^0, y^1) = \hat{\sigma}_i(y^0)(y^1)$, etc.

This is a PPE with payoffs $v$, so $E(\delta) \supseteq B(\delta, E(\delta))$.

**Definition** (APS) $W$ is *self generating* if $W \subseteq B(\delta, W)$

$W$ is self generating if the set of payoffs that can be enforced with continuation payoffs in $W$ includes all of the payoffs in $W$. 
A trivial example of a self-generating set is the point corresponding to the payoffs at a particular static equilibrium.

Are there any other one-point self-generating sets?
Theorem (APS):
If $W$ is self generating and bounded, then $W \subseteq E(\delta)$

Proof: Fix a $v \in W$. Construct strategies that (a) yield $v$ and (b) are PPE. First construct the strategies:

**Step 1:** $v \in B(\delta, W)$ (why?) so there is $\alpha^0$ and a $w^0 : Y \to W$ that generate $v$. Set date-0 actions to be $\sigma^0 = \alpha^0$

**Step 2:** for each $y^0$, set $v^1(y^0) = w^0(y^0)$. Every $v^1 \in B(\delta, W)$, so there are $\alpha^1(v^1), w^1(y^1, v^1)$ that generate $v^1$. Set $\sigma^1(y^0) = \alpha^1(v^1) = \alpha^1(w^0(y^0))$. And set $v^2(y^0, y^1) = w^1(y^1, v^1(y^0))$.

Now define date-2 strategies, etc.
Claim 1: These strategies yield payoff $v$.

This uses the boundedness of $W$. To see why, consider a 1-player game where all actions yield payoff 0. Then since 0 is only feasible payoff, it is only equilibrium payoff, and the incentive constraint is moot, but the whole real line is self generating, i.e. $R \subseteq B(\delta, R)$!

(To see why, take any $v \in R$, set $w(y) = v / \delta$.)
What goes wrong with the proof? We can still construct the strategy in the same way.

The construction runs as follows: stage-0 payoff is 0, so the continuation payoff $w^0$ is $v/\delta$. Stage 1 payoff is again 0, payoffs from period 1 on are supposed to add up to $v/\delta$, so make the continuation payoffs from period 2 on be $v/\delta^2$.

Etc. At each date we can generate the needed payoff by a combination of 0 today and even more in the future.

To prove claim 1, use compactness of payoff set to argue that the average present value of our sequence is about the same as if we truncated at time $T$ and had payoff 0 from then on (“continuity at infinity”), and take a limit on the truncation point.
Claim 2: The strategy profile constructed is a PPE.

This is from continuity at infinity and the 1-stage deviation principle, which they satisfy by construction. (FT chapter 4).
Simple example of a self-generating set illustrates the definitions.
Start with a prisoner’s dilemma

<table>
<thead>
<tr>
<th></th>
<th>C</th>
<th>D</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td>2,2</td>
<td>−1,3</td>
</tr>
<tr>
<td>D</td>
<td>3,−1</td>
<td>0,0</td>
</tr>
</tbody>
</table>

Observed actions: there are 4 public outcomes corresponding to the 4 action profiles.

Let $W = \{v', v''\}$, where

$v' = \left(\frac{3-\delta}{1+\delta}, \frac{3\delta-1}{1+\delta}\right)$ and $v'' = \left(\frac{3\delta-1}{1+\delta}, \frac{3-\delta}{1+\delta}\right)$
Claim: W is self-generating for $\delta \geq 1/3$.

To prove this we need to generate both $v'$ and $v''$ with continuation payoff is $W$.

To generate $v'$ set $\alpha = (D, C)$

$$w(D, C) = w(C, C) = v''$$
$$w(D, D) = w(C, D) = v'$$

Note that the continuation payoffs implicitly define a “transition function” for the “state:”
Namely, if 2 plays D- stick at $v'$, else got to $v''$.

Check that this adds up to the right payoff vector:
\[(1 - \delta)(3, -1) + \delta v'' =
\]
\[
\left( \frac{3(1 - \delta^2) + 3\delta^2 - \delta}{1 + \delta}, \frac{-(1 - \delta^2) + 3\delta - \delta^2}{1 + \delta} \right) = v'
\]

Now check IC constraints.

The continuation payoff is independent of 1's action so 1 will play a SR best response, so 1 plays D.

2: told to play C.

conform: \[v'_{2} = \frac{3\delta - 1}{1 + \delta}\]

deviate to D: \[0 + \delta v'_2\], so 2 conforms if \[v'_2 \geq 0\] or \[\delta \geq 1/3\].
Properties of $E(\delta)$.

- $E(\delta)$ is compact.

(This is a consequence of more general results on dynamic games, see FL JET 83- basically we can use the product topology on the strategy space.)

- $E(\delta)$ needn’t be monotone in $\delta$, and it needn’t be convex, but roughly speaking “convexity implies monotonicity.”

*Intuition for Monotonicity:* Increasing $\delta$ makes the future more important, so we expect it to lead to “more” equilibria and a larger set of equilibrium payoffs. In particular, as $\delta$ increases, “less variation” in continuation payoffs is needed to enforce a given profile $\alpha$. And with convexity, these less variable continuation payoffs will also be equilibria.
But the set of equilibrium payoffs is not in general monotone (in the sense of set inclusion) in $\delta$.

**Reason**: $E(\delta)$ needn’t be convex.

Lack of convexity easiest to see for small $\delta$.

With convexity, things are nicer:

For any set $X$ let $co(X)$ denote the convex hull of $X$. 
Lemma: Suppose $0 < \delta_1 < \delta_2$, $W \subseteq W'$, and $W \subseteq B(\delta_1, W')$. Then $W \subseteq B(\delta_2, co(W'))$.

Proof: Fix $v \in W \subseteq B(\delta_1, W')$.
Suppose $(\alpha, w)$ generates payoff $v$ w.r.t $(\delta_1, W')$.
We will find $w'$ such that $(\alpha, w')$ generates payoff $v$ w.r.t $(\delta_2, co(W'))$.

Define $w'$ by $w'(y) = \frac{(\delta_2 - \delta_1)}{(\delta_2(1 - \delta_1))}v + \frac{\delta_1(1 - \delta_2)}{\delta_2(1 - \delta_1)}w(y)$.

This has “less variation” than $w$ because $\delta_2 > \delta_1$.

Since $v$ and $w$ are in $W'$, the $w'(y) \in co(W')$.

And can check $w'$ enforces $\alpha, v$ when discount factor is $\delta_2$: 
Recall: 
\[ w'(y) = \frac{(\delta_2 - \delta_1)}{(\delta_2(1 - \delta_1))} v + \frac{\delta_1(1 - \delta_2)}{\delta_2(1 - \delta_1)} w(y) \]

So

\[
(1 - \delta_2)g_i(a_i, \alpha_{-i}) + \delta_2 \sum_y \pi_y(a_i, \alpha_{-i})w'(y) =
\]

\[
\frac{\delta_2 - \delta_1}{1 - \delta_1} v + \frac{1 - \delta_2}{1 - \delta_1} \left\{ (1 - \delta_1)g_i(a_i, \alpha_{-i}) + \delta_1 \sum_y \pi_y(a_i, \alpha_{-i})w(y) \right\}
\]

Term inside curly brackets is the payoff to \( a_i \) when discount factor is \( \delta_1 \) and continuation is \( w \). So \( i \) doesn’t want to deviate, and overall payoff is \( v \).
So we have proved the lemma:

If \( 0 < \delta_1 < \delta_2, W \subseteq W', \) and \( W \subseteq B(\delta_1, W'), \) then \( W \subseteq B(\delta_2, co(W')). \)

**Corollary:** If \( E(\delta_1) \) is convex, then for any \( \delta_2 \in (\delta_1, 1), E(\delta_1) \subseteq E(\delta_2). \)

**Proof:** Set \( W = E(\delta_1) = W'. \) Then \( W = E(\delta_1) \subseteq B(\delta_1, W') \) and so \( E(\delta_1) \subseteq B(\delta_2, E(\delta_1)): \) the set \( E(\delta_1) \) is self-generating under the higher discount factor \( \delta_2. \)
One case where $E(\delta)$ is convex is when players observe the outcome of a continuously-distributed public randomizing device at the start of each period. (these signals are iid.)

Then we can prove monotonicity directly, working with the strategies instead of the continuation values.

We expect $E(\delta)$ to become convex as the discount factor goes to 1, since deterministic cycles can replicate the effect of public randomizations.

And we will show that the set of “limit PPE payoffs” $\lim_{\delta \to 1} E(\delta)$ is convex, and the same with and w/o public randomizations.

So might think this is true before the limit- but Yamamoto [2009] has an example where both convexity and monotonicity fail for all $\delta$. 
Another case where the set of PPE is convex is when the set of public signals lies in a continuum and has a continuous density, as in APS- provided that the agents observe one draw of the signal process before taking any actions.

If the first observation of \( y \) comes after first period play then the set of PPE needn’t be convex (think of \( \delta = 0 \)) but the set of continuation PPE is.

**Theorem** (APS 1990) Suppose that signals have a continuous density on \( \mathbb{R}^k \), with support independent of the action profile, that payoffs are continuous in the signal, and let \( W \) be compact. Then any pure action \( a \) that is enforceable wrt conhull (W) can be enforced with continuation payoffs that take values in \( \text{ext}(W) \).

Here “ext(W)” is the extreme points: the points that aren’t convex combinations of points in conhull (W). The conclusion is trivial with a public randomizing device at the start of each period, much harder w/o it.
**Intuition:** a continuously-distributed (non-atomic) signal lets us construct any finite-support public randomizing device by “subdividing the interval” into small pieces whose overall probability is unaffected by the player’s actions.

Easy to do this and match the distribution of signals from a single action profile $a$, the proof shows we can do this for any finite set of actions (and so keep the IC constraints satisfied.)
Very rough sketch of proof: Look at the set of continuation payoff functions that enforce action $a$ wrt continuation payoffs in conhull $(W)$ and has a given overall utility. This set is convex and compact (in the weak-* topology) and so (because conditions of the Krein-Milman theorem are satisfied) it is the convex hull of its extreme points. Then because outcomes have non-atomic measure, and finitely many incentive constraints, show that the extremal function $w$ takes on extremal values in $W$. (not true if e.g. a single outcome under action $a$.) (Lyapunov’s convexity theorem- easy proof in Ross, Am. Math Monthly 2005, 651-652.)

Let $B^p(W, \delta)$ be the set of payoffs that can be generated by pure actions.

Corollaries: Under the assumptions of APS, if $W$ is compact, then $B^p(W, \delta) = B^p(conhull(W), \delta)$, so the set of payoffs to pure-strategy equilibria is (weakly) monotonic in the discount factor.
In principle could use the bang-bang result to characterize the PPE payoffs for a fixed $\delta$ but this is hard to apply in practice. *(seldom if ever been done?)*

It is much easier to compute the set of “strongly symmetric” PPE of a symmetric game.

Strong symmetry: $\alpha_i(h_t) = \alpha_j(h_t)$ for all histories $h_t$, even those that are asymmetric.

In repeated prisoner’s dilemma with observed actions, “Tit for Tat” is symmetric but not strongly symmetric.

Computing the set of strongly symmetric PPE boils down to computing two numbers, the highest and lowest equilibrium payoff, using the fact that these payoffs can be generated by continuation payoffs that themselves only take on these values.
The Impact of Increased Precision

Kandori (1992) shows that improving the precision of the public signals cannot reduce the set of equilibrium payoffs.

To rank the informativeness of the public signals, use Blackwell’s (1951) idea of a garbling:

View the realized signal $y$ as the result of an experiment about the action profile $a$. Two different public monitoring distributions (with different signal spaces) can then be viewed as two different experiments.

Let $\Pi$ denote the $A \times Y$ matrix whose $a$th-row corresponds to the probability distribution over $Y$ under the action profile $a$. 
**Definition** The public monitoring structure $\Pi'$ on signal space $Y'$ is a garbling of monitoring structure $\Pi$ on $Y$ if there exists a stochastic matrix $Q$ such that $\Pi' = \Pi Q$.

That is, $\pi'_{y'}(a) = \sum_y \pi_y(a) \cdot q(y'|y)$

(Stochastic matrix: non-negative, rows sum to 1.)

$Y'$ may have more or less elements than $Y$.

Kandori calls this a quasi-garbling.

Garbling is a partial order, not an order, and it is not strict. For example, if $Q$ is a permutation matrix, then $Y'$ is simply a re-labeling of $Y$, and $\Pi$ and $\Pi'$ are garblings of each other.
If public randomizations are available, and \( \Pi' \) is a garbling of \( \Pi \), then any equilibrium under \( \Pi' \) can be replicated under \( \Pi \) by using the public randomizing device to do the garbling and then playing the equilibrium for \( \Pi' \).

Even if there aren’t public randomizations we have the following result.

**Proposition:** If \( \Pi' \) is a garbling of \( \Pi \), and \( W \subseteq B(\delta, W', \Pi') \), then \( W \subseteq B(\delta, \text{conhull}(W'), \Pi) \).

So \( E_{\Pi'}(\delta) \subseteq E_{\Pi}(\delta) \) if the set of equilibrium payoffs is convex.

The proof is like the proof of monotonicity in the discount factor… On stronger assumptions- strictly positive Q plus “Slater condition”—Kandori shows strict monotonicity of pure equilibria in APS setting.
Next step: characterize the set of PPE payoffs when players are patient, i.e. determine \( \lim_{\delta \to 1} E(\delta) \).

Typical question of interest: are there efficient or approximately efficient PPE?

Possible motivations:
  a) *We might* think that players will tend to play efficient equilibrium when these are possible, and
  b) players may look harder for ways to change the rules of the game if the game they are in keeps them from getting good payoffs.

The characterization will rely on an LP algorithm- details next class.
First a little intuition:

1) Ask how far the set of PPE payoffs can extend in each direction.

The PPE payoffs can’t go farther than feasible set- no PPE can give a higher sum of payoffs than the highest feasible sum. When does the PPE set reach all the way to the frontier?

For example, pick direction $45^\circ$, and a point that gives the best feasible payoff vector in that direction- that is, a point $v^*$ that maximizes the sum $v_1 + v_2$ over the feasible set $V = \text{conhull}\{v \mid \exists a \text{ s.t. } g(a) = v\}$

Let $a^*$ be the corresponding action profile, and let the maximum sum be $k^*$.

Suppose that there is a single action profile that is maximal in this direction, so $v^*$ is at a kink in the frontier of the feasible set.
Try to generate payoff vector \( v^* \) using first period actions \( a^* \), continuation payoff function \( w \) whose values must be equilibrium payoff vectors.

For each \( y \) the continuation payoff \( w(y) \) can’t have a higher "score" (sum of payoffs) than \( k^* \). (why?)

Moreover, if the expected continuation score is lower than \( k^* \), the overall score has to be lower than \( k^* \).

Thus the only way to generate a sum of payoffs equal to \( k^* \) is if for every outcome \( y \) that has positive probability, the continuation payoff is on the line \( w_1(y) + w_2(y) = k^* \).

When can this happen? Examples:
a) if \( a^* \) is static equilibrium, or
b) if observed actions.
c) more generally than b) it enough that all “tempting deviations” be observed.

Conversely, if every outcome has positive probability under \( a^* \) and \( a^* \) is not a static equilibrium, then in order to satisfy the IC constraints there has to be positive probability of at least 2 different continuation payoffs.

And since \( k^* \) is at a kink, one of these continuation payoffs must have a sum of payoffs that is less than \( k^* \).

**Conclusion:** In general, we shouldn’t expect to be able to generate a score of exactly \( k^* \) for any fixed discount factor, so there may be no equilibrium with sum of payoffs equal \( k^* \).

But there may be equilibria that are approximate this sum, where the approximation error vanishes as players become patient.
As $\delta \rightarrow 1$, a given payoff vector $v$ can be generated with less variation in the continuation payoffs, because the future is more important compared to the present.

In some cases this will permit the equilibrium payoffs to converge to efficiency. A sufficient condition is that the efficient payoff in question can be enforced with continuation payoffs on a hyperplane parallel to the tangent to the equilibrium set.

*Intuition:* Here the “efficiency loss” relative to the feasible set vanishes as $\delta \rightarrow 1$.

But the efficiency loss *doesn’t* vanish when the continuation payoffs are orthogonal to the tangent.
Next time: more on repeated games with imperfect public monitoring
FL [1994]
FLM [1994]
Athey-Bagwell [2001]
Levin [2003]