The Prisoners' Dilemma Again (and Again)

Recall (a version of) the Prisoners' Dilemma game. Wouldn't it be nice if players could cooperate?

	C	D
C	$3, 3$	$0, 5$
D	$5, 0$	$1, 1$

Perhaps if the game is played repeatedly, cooperation would be possible in early periods by threatening to defect later if cooperation was not observed? Suppose the game is repeated T times:

- In the last time period (T), D is a dominant strategy for both players.
- In period $T - 1$ neither player can influence future decisions so each will play D.
- This logic continues: iterating back to the first period, both players play D throughout.

Thus the unique subgame-perfect equilibrium involves defection in every period, regardless of previous play. This is also the unique Nash equilibrium! Why?

...and Again

Consider a strategy that called upon a player to play C in some period $1 \leq t \leq T$. This is dominated by a strategy that is identical everywhere except that it calls for D at period t.

Continuing (and tightening) this iterated dominance argument yields a unique Nash equilibrium.

However, this depends upon the *uniqueness* of the equilibrium in the one-shot game...repeating a game in this manner will usually expand the set of achievable outcomes in the *stage game*.

	L		M		R	
T	0	0	3	4	6	0
M	4	3	0	0	0	0
B	0	6	0	0	5	5

In the above game $\{B, R\}$ is not a Nash equilibrium of the stage game, but it can be played in the first period, as part of a subgame-perfect equilibrium, if the game is played twice.

Repeated Games with Discounting

A stage game is a (mixed extension of the) strategic-form game $\mathcal{G} = \langle N, \{S_i\}_{i \in N}, \{u_i\}_{i \in N} \rangle$.

Definition 28. The repeated game with discounting of the stage-game $\mathcal{G}$ is the extensive-form game $\Gamma = \langle N, H, P, \{U_i\}_{i \in N} \rangle$ with,

1. $H = \cup_{t=0}^{T} S^t$ (where $S^0 = \emptyset$ is the initial history and T is the number of *stages*).
2. $P(h) = N$ for each non-terminal history $h \in H$.
3. Payoffs involve a discount factor, $\delta \in (0, 1)$, and are such that

$$U_i = \sum_{t=1}^{T} \delta^{t-1} u_i(s^t),$$

where $s^t \in S$ is the strategy profile of the game $\mathcal{G}$ played in stage $t \in \{1, \ldots, T\}$.

For infinitely repeated games let $T = \infty$ and (normalising to per-period payoffs)

$$U_i = (1 - \delta) \sum_{t=1}^{\infty} \delta^{t-1} u_i(s^t).$$

The Stage Game

So: how do players play $\{B, R\}$ in equilibrium in the following game when repeated twice?

	L		M		R	
T	0	0	3	4	6	0
M	4	3	0	0	0	0
B	0	6	0	0	5	5

The pure-strategy Nash equilibria of the stage game are $\{T, M\}$ and $\{M, L\}$.

Consider a mixed-strategy σ for row player placing probability p on T and $1 - p$ on M. Row gets:

	L	M	R
σ	$4(1-p)$	$3p$	$6p$
B	0	0	5

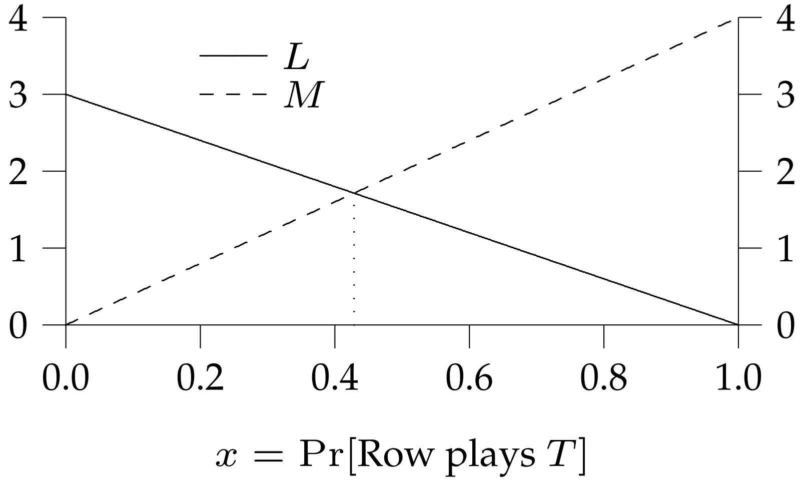
This shows row's payoffs from the mixed-strategy σ and from the pure-strategy B. Thus σ dominates B, with $1 > p > \frac{5}{6}$. By symmetry, R is dominated for column player.

The Reduced Stage Game

Deleting strictly dominated strategies yields the following reduced game:

	$L\ (y)$	$M\ (1-y)$	Expected
$T\ (x)$	0 0	3 4	$3(1-y)$
$M\ (1-x)$	4 3	0 0	$4y$
Expected	$3(1-x)$	$4x$	

Note that there are three Nash equilibria of this reduced game, two pure and one mixed:



$$\Rightarrow \quad Z = \begin{cases} \{T, M\} & \Rightarrow u = (3, 4) \\ \{M, L\} & \Rightarrow u = (4, 3) \\ \{\frac{3}{7}, \frac{3}{7}\} & \Rightarrow u = (\frac{12}{7}, \frac{12}{7}) \end{cases}$$

The strategy profile $\{B, R\}$ Pareto dominates each of these equilibria. How can this be obtained?

Repeating the Stage Game: Conditional Strategies

The game is played twice, with discount rate $\delta < 1$. Suppose $\{B, R\}$ is played in the first period.

	L		M		R	
T	0	0	3	4	6	0
M	4	3	0	0	0	0
B	0	6	0	0	5	5

$$\Rightarrow \quad \text{with } Z = \begin{cases} \{T, M\} & \Rightarrow u = (3, 4) \\ \{M, L\} & \Rightarrow u = (4, 3) \\ \left\{\frac{3}{7}, \frac{3}{7}\right\} & \Rightarrow u = \left(\frac{12}{7}, \frac{12}{7}\right) \end{cases}$$

Conditional Strategy: if $\{B, R\}$ is observed in the first period, row plays M and column plays L.

If anything else is observed, row player places probability $\frac{3}{7}$ on T and $\frac{4}{7}$ on M. Column player places probability $\frac{3}{7}$ on L and $\frac{4}{7}$ on M. These strategies constitute a subgame-perfect equilibrium.

- In the second period (nine subgames) play corresponds to a Nash equilibrium.
- In the first period (a single subgame), by following the above strategy, row gets $5 + 4\delta$.
- Column gets $5 + 3\delta$. By deviating, the greatest column could get would be $6 + \frac{12}{7}\delta$.

This is a (subgame-perfect) equilibrium so long as: $\quad 5 + 3\delta \geq 6 + \frac{12}{7}\delta \quad$ or: $\quad \delta \geq \frac{7}{9}$.

Another Example

Consider the below game, played twice with a discount factor $\delta < 1$.

	L		M		R	
T	2	2	0	0	6	0
M	0	0	4	4	0	0
B	0	6	0	0	5	5

Note that $\{B, R\}$ is not a Nash equilibrium of the stage game. Nevertheless, consider:

- Play $\{B, R\}$ in the first period. If $\{B, R\}$ observed, play $\{M, M\}$ in the second.
- If $\{B, R\}$ is not observed in the first period, play $\{T, L\}$.
- Payoff from playing strategy is $5 + 4\delta$.
- Payoff from deviating is at most $6 + 2\delta$.

If $\delta \geq \frac{1}{2}$ these strategy profiles are a subgame-perfect equilibrium.

Equilibria in Repeated Games

Suppose this game is repeated three times. How many subgames are there? Showing that a given strategy profile is a subgame-perfect equilibrium might seem like a daunting task. However…

The **One-Deviation Principle** will simplify this task.

Additionally in repeated games, the number of subgame-perfect equilibria might be very large indeed. Characterising them would once again seem like a *very* daunting task. However…

The **Folk Theorems** will simplify this task.

The rest of the lecture introduces these ideas and applies them to some examples.

The One-Deviation Principle

Definition 29. A strategy for player i satisfies the one-deviation principle (or property) if for any history $h \in H$ such that $i \in P(h)$, there is no deviation that i could make to increase their payoff whilst leaving all the other players' strategies fixed, *and the rest of their own strategy*.

> "A strategy profile in a finite-horizon extensive-form game or in an infinitely repeated game with discount factor $\delta < 1$ is a subgame-perfect equilibrium if and only each player's strategy satisfies the one-deviation principle."

- In other words, only need to check one-deviation-at-a-time at every stage for each player.
- Ignore multiple contemporaneous deviations or multiple sequential deviations.

	L	M	R
T	2 2	0 0	6 0
M	0 0	4 4	0 0
B	0 6	0 0	5 5

$(\times\ 3)\ldots$

Thrice-Repeated Example

1. Consider the strategy: Play B, then B if $\{B, R\}$ is observed and T if anything else, then M if $(\{B, R\}, \{B, R\})$ is observed and T if anything else (for row player).

2. Suppose column player is playing R, then R if $\{B, R\}$ is observed and L if anything else, then M if $(\{B, R\}, \{B, R\})$ is observed and L if anything else.

3. Need only check single deviations for each player at each stage. No profitable deviations in the last stage — if $\{B, R\}$, $\{B, R\}$ is observed M is a best response, and if not, T is a best response.

4. At penultimate stage, B yields a higher payoff if $5 + 4\delta \geq 6 + 2\delta$ (or $\delta \geq \frac{1}{2}$). This second payoff arises from the most profitable single deviation possible at this stage.

5. At initial stage, B yields a higher payoff if $5 + 5\delta + 4\delta^2 \geq 6 + 2\delta + 2\delta^2$ (or $\delta \geq 0.28$ish). This second payoff arises from the most profitable single deviation possible at this stage.

6. Symmetry implies column player's strategy is also subgame perfect for $\delta \geq \frac{1}{2}$.

No need to check any other deviant strategies: e.g. Play B *then* T...

Multiple Equilibria in Repeated Games

There are many such subgame-perfect equilibria in this repeated game. In fact, when games are repeated, many many different outcomes can be supported as equilibria. e.g.

- Playing the same Nash equilibrium in every stage is always subgame perfect.
- Playing any sequential combination of Nash equilibria is subgame perfect.
- Conditioning the future Nash equilibrium to be played on current choices…
- …allows non-Nash strategies to be part of subgame-perfect equilibria.

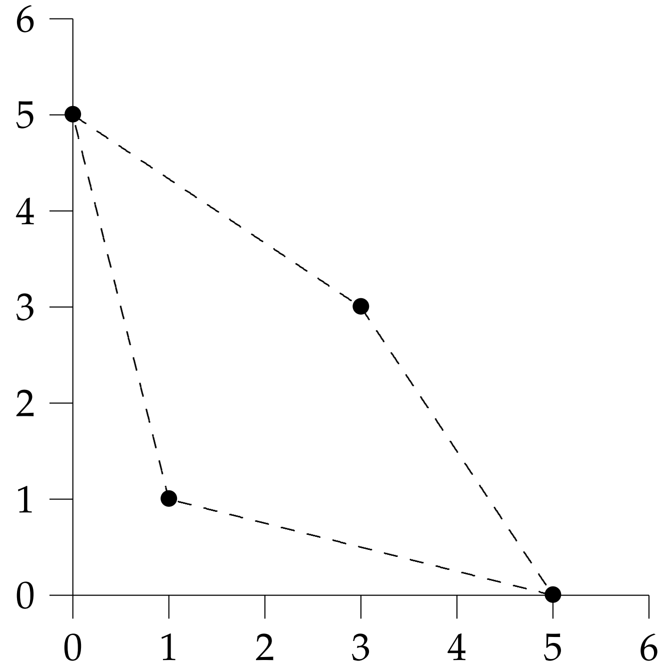
Rather than characterise all these equilibria, characterise the (normalised one-period) *payoffs* that are achievable as part of a Nash (and subgame-perfect) equilibrium.

- In the finitely repeated Prisoners' Dilemma, this argument did not work…
- …because there is a *unique* Nash equilibrium of the stage game and hence…
- …there is no choice of equilibria with which to condition behaviour.
- In the infinitely repeated game, however, there is a multiplicity of equilibria.

Characterising a Game in Payoff Space

Consider the Prisoners' Dilemma. Plotting row player's payoffs against column player's payoffs, and allowing players to mix, which payoffs are achievable in the one-shot game?

	C	D
C	3, 3	0, 5
D	5, 0	1, 1



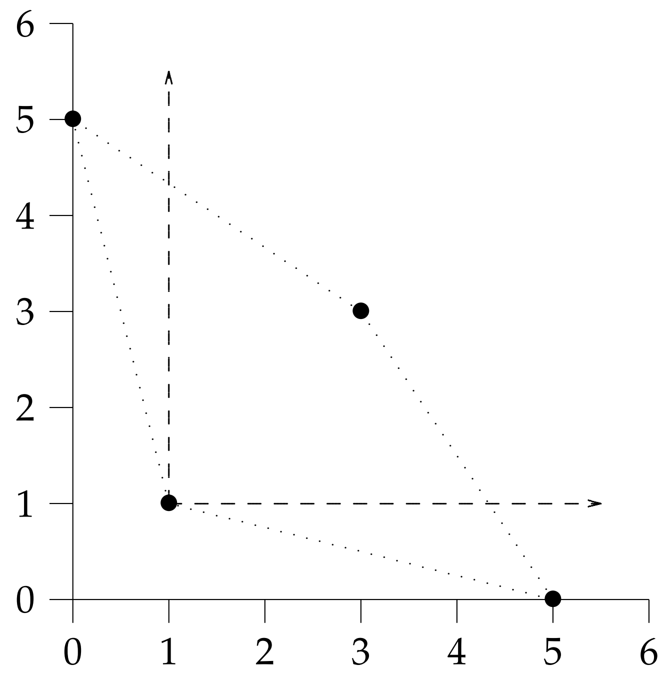
Appropriate mixtures over the two strategies generate all the payoffs inside the diamond. This is the *convex hull* of the payoffs to pure strategies. All these payoffs are *feasible*.

Which of these are supportable as part of an equilibrium in an infinitely-repeated game?

The Nash-Threats Folk Theorem

"Every feasible payoff profile above the Nash equilibrium payoff profile can be achieved by a subgame-perfect equilibrium of the infinitely-repeated game for δ large enough."

	C	D
C	3, 3	0, $\underline{5}$
D	$\underline{5}$, 0	$\underline{1}$, $\underline{1}$



How? Play strategies that generate required payoff combination. If anyone deviates, play the stage Nash equilibrium forever. Consider payoff profile $(3, 3)$ in the Prisoners' Dilemma. . .

Cooperating in the Prisoners' Dilemma

Compare the payoff stream from $\{C, C\}$ forever with the payoff stream from a single deviation in any subgame (using one-deviation principle and noting all subgames look the same):

$$3 + 3\delta + 3\delta^2 + \ldots = \frac{3}{1-\delta} \geq 4 + \frac{1}{1-\delta} = 5 + \delta + \delta^2 + \ldots$$

This requires $\delta \geq \frac{1}{2}$. The "grim" strategies (play C forever unless D is ever observed, in which case play D forever) constitute a subgame-perfect equilibrium. Logic carries over to collusion…

Consider an infinitely-repeated n-firm Bertrand pricing game. Charge the monopoly price with profits π_M/n. Any deviation prompts marginal-cost pricing for T periods.

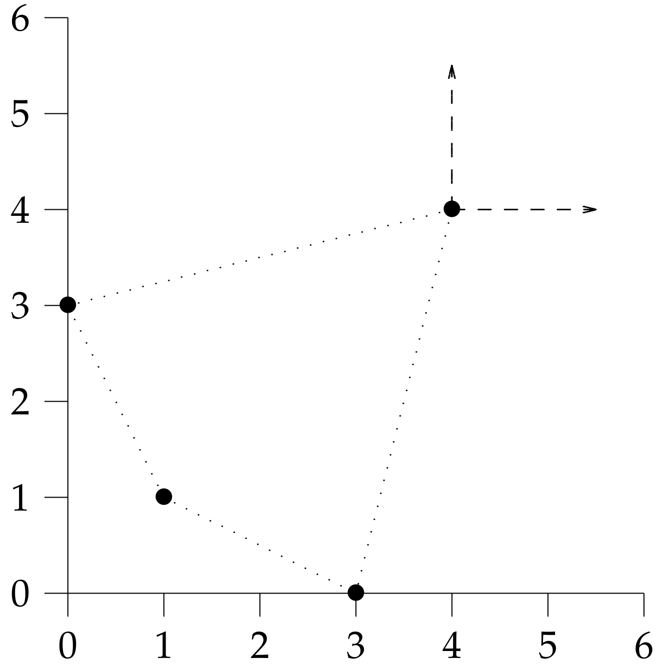
$$\frac{\pi_M}{n}\sum_{t=1}^{\infty}\delta^{t-1} = \frac{1}{1-\delta}\frac{\pi_M}{n} \quad \text{versus} \quad \pi_M + \frac{\delta^{T+1}}{1-\delta}\frac{\pi_M}{n}.$$

The former is greater than the latter if $1 - \delta^{T+1} \geq n(1-\delta)$, which holds certainly if δ and T are large enough (in the grim strategy, T is infinite, and $\delta \geq 1 - \frac{1}{n}$ is the appropriate condition).

Even More Equilibria

Consider the following game, with its associated payoff representation.

	C	D
C	$\underline{4}$, $\underline{4}$	$\underline{3}$, 0
D	0, $\underline{3}$	1, 1



The Nash-threat Folk Theorem only indicates that $(4, 4)$ can be achieved. But many other payoffs can be achieved by subgame-perfect equilibria also. To do this, need to define “minmax” payoffs:

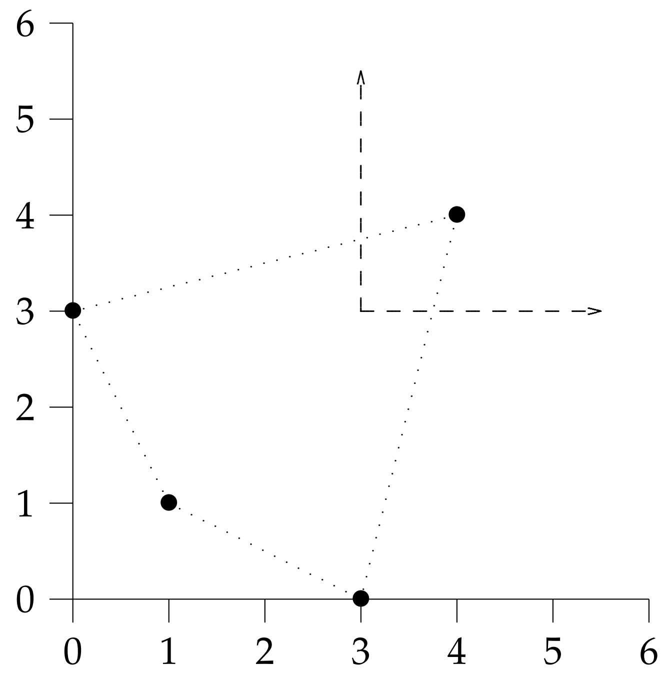
The lowest payoff one player can force the other player to, *given* the other player best-responds.

Minmax Payoffs and another Folk Theorem

Player i's minmax payoff is given by: $u_i^m = \min_{s_{-i}} \left\{ \max_{s_i} u_i(s) \right\}$. The folk theorem says:

> "Every feasible payoff profile above the minmax payoff profile can be achieved by a subgame-perfect equilibrium of the infinitely-repeated game for δ large enough." (There is a technical *full dimensionality* condition.)

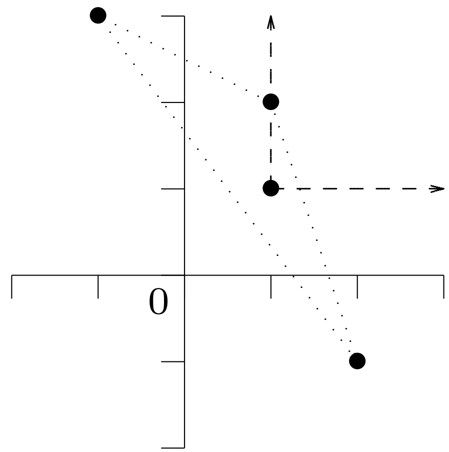
	C	D
C	$\underline{4}$, $\underline{4}$	$\underline{3}$, 0
D	0, $\underline{3}$	1, 1



Another Example

Consider the following game. What are the minmax payoffs? What are the subgame-perfect equilibrium payoffs achievable in the infinitely-repeated version?

	L	R
T	-1 3	2 -1
M	2 -1	-1 3
B	1 2	1 2



Minmax payoffs are the worst thing that others can do to a player, whilst that player best responds.

- Minmax payoff for the row player is 1, e.g. by column playing $L\&R$ with probability $\frac{1}{2}$ each.
- Minmax for the column player is 1, e.g. play T with probability $\frac{1}{2}$ and M with probability $\frac{1}{2}$.

So: every feasible payoff combination above $(1, 1)$ is the outcome of a subgame-perfect equilibrium of the infinitely-repeated game. Requires complex "punishment and reward" strategies.

Folk Theorems with Incomplete Information

Even when an opponent's actions are not directly observable, similar results are available. The following example illustrates (formal statements are beyond the scope of the course):

- Recall the repeated Bertrand model introduced briefly earlier.
- Suppose now that the price set by one firm is not observable by the other.
- Strategies are conditioned on some publicly-observed state variable, e.g. market demand.

Suppose that with probability $(1-\alpha)$, demand follows from the standard Bertrand model, but with probability α, demand falls to zero (owing to conditions outside the firms' control). Strategies:

- Play p_M (monopoly price) and split the market as long demand is non-zero.
- Play p_C (marginal cost) for T periods if zero demand is observed by either firm.
- Note the firm that cheated (set a price $p < p_M$) will know that the other firm has observed zero demand in that period! Hence both firms know when a "price war" will start.

This last feature is critical to the ensuing analysis. First, calculate payoffs…

Calculating Payoffs

Start by considering the payoffs to a player that accrue when players collude and charge p_M. Denote expected payoffs in the collusive phase and in a price war phase as V_C and V_W respectively. Then,

$$V_C = \underbrace{(1-\alpha)\left[\frac{\pi_M}{2} + \delta V_C\right]}_{\text{High Demand}\Rightarrow\text{Retain C Phase}} + \underbrace{\alpha\delta V_W}_{\text{Enter War}} \qquad \text{and} \qquad V_W = \underbrace{\delta^T V_C}_{\text{Wait T Periods for C Phase}},$$

where π_M is monopoly profit. A deviating player would obtain at most V_D, where

$$V_D = (1-\alpha)\left[\pi_M + \delta V_W\right] + \alpha\delta V_W.$$

So if the following inequality obtains, these strategies will form an equilibrium:

$$(1-\alpha)\left[\pi_M + \delta V_W\right] + \alpha\delta V_W \leq (1-\alpha)\left[\frac{\pi_M}{2} + \delta V_C\right] + \alpha\delta V_W.$$

Supporting Collusion

Subtracting $\alpha\delta V_W$ from both sides, and dividing by $(1-\alpha)$, this condition becomes

$$\pi_M + \delta V_W \leq \frac{\pi_M}{2} + \delta V_C \qquad \text{or} \qquad \frac{\pi_M}{2} \leq \delta(V_C - V_W).$$

Solving simultaneously for V_C and V_W from the initial equalities on the previous slide yields

$$V_C = \frac{(1-\alpha)\pi_M/2}{1-(1-\alpha)\delta - \alpha\delta^{T+1}} \qquad \text{and} \qquad V_W = \frac{(1-\alpha)\delta^T \pi_M/2}{1-(1-\alpha)\delta - \alpha\delta^{T+1}}.$$

Finally, substituting into the inequality above gives the condition

$$2(1-\alpha)\delta - (1-2\alpha)\delta^{T+1} \geq 1.$$

This is satisfied for appropriate values of δ and α given a large enough T, (i.e. for α small, δ large and T large). So Folk Theorems are available under incomplete information.