

Extensive Form Games

Let's begin informally.

Consider the following entry game between an entrant and an incumbent

IMAGE

What should we expect players to do?

Let's begin with the concept we already have: Nash equilibrium.

Let's use the normal form of the game:

IMAGE

There are two Nash equilibria:

No entry, Fight if entry

Entry, Accommodate if entry

But there is something weird about the first one:

The incumbent is not willing to Fight if the Entrant actually enters the market

This is a non-credible threat

To use this idea, we must use the time structure of the game, which is absent in the normal form, which is most useful exactly for static games

Methodologically, we'll solve the game backwards

Remark on commitment

What matters is commitment power

If the Incumbent could commit to Fight, then (No entry, Fight) would indeed be a good forecast for the outcome of this game

The time structure allows for commitment power: once it's done, it's done – in our example, the entry by the entrant cannot be undone

Strategies in dynamic games

Consider the following game:

IMAGE

What are the strategies for players 1 and 2?

Main idea: a strategy must determine an action for each possible contingency in which the game might reach.

The player could write down all his decisions, for all possible contingencies, give it to a computer to implement, and walk away.

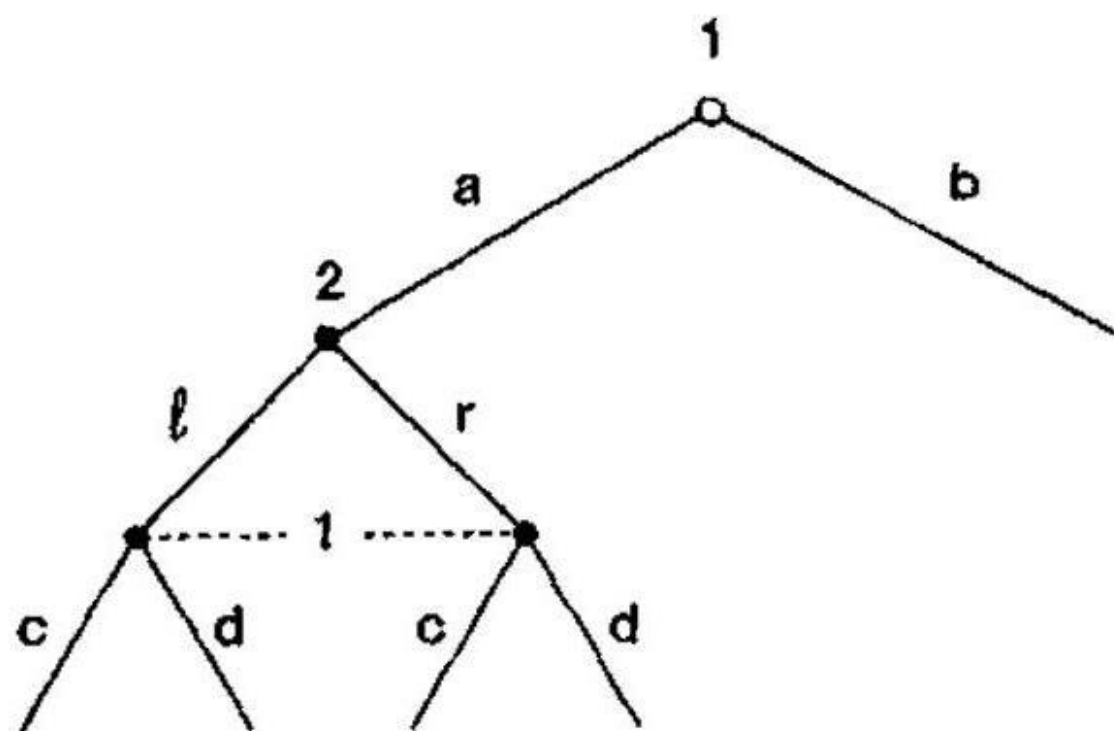
This holds even if one action makes it impossible for some contingency to be achieved.

IMAGE

What happens off the equilibrium path matters!

Extensive-Form Games: main ingredients

- N : finite set of players; nature is player $0 \in N$
- tree: order of moves
- payoffs for every player at the terminal nodes
- information partition
- actions available at every information set
- description of how actions lead to progress in the tree
- random moves by nature



Extensive-Form Games: full description

1. Game Tree

- $(X, >)$: tree
- X : set of nodes
- $x > y$: node x precedes node y
- $\phi \in X$: initial node, $\phi > x, \forall x \in X \setminus \{\phi\}$
- $>$ transitive ($x > y, y > z \Rightarrow x > z$) and asymmetric ($x > y \Rightarrow y \not> x$)
- every node $x \in X \setminus \{\phi\}$ has one immediate predecessor: $\exists x' > x$ s.t. $x'' > x \& x'' \neq x' \Rightarrow x'' > x'$
- $Z = \{z \mid \nexists x, z > x\}$: set of terminal nodes
- $z \in Z$ determines a unique path of moves through the tree, payoff $u_i(z)$ for player i

2. Information Partition

- information partition: a partition of $X \setminus Z$
- node x belongs to information set $h(x)$
- player $i(h) \in N$ moves at every node x in information set h
- $i(h)$ knows that he is at some node of h but does not know which one
- same player moves at all $x \in h$, otherwise players might disagree on whose turn it is

$$i(x) := i(h(x))$$

3. Actions

- $A(x)$: set of available actions at $x \in X \setminus Z$ for player $i(x)$
- $A(x) = A(x') =: A(h), \forall x' \in h(x)$ (otherwise $i(h)$ might play an infeasible action)
- each node $x \neq \phi$ associated with the last action taken to reach it
- every immediate successor of x labeled with a different $a \in A(x)$ and vice versa
- move by nature at node x : probability distribution over $A(x)$

4. Strategies

- $H_i = \{h \mid i(h) = i\}$
- $S_i = \prod_{h \in H_i} A(h)$: set of pure strategies for player i
- $s_i(h)$: action taken by player i at information set $h \in H_i$ under $s_i \in S_i$
- $S = \prod_{i \in N} S_i$: strategy profiles
- A strategy is a complete contingent plan specifying the action to be taken at each information set.
- Mixed strategies: $\sigma_i \in \Delta(S_i)$
- mixed strategy profile $\sigma \in \prod_{i \in N} \Delta(S_i) \rightarrow$ probability distribution $O(\sigma) \in \Delta(Z)$
- $u_i(\sigma) = \mathbb{E}_{O(\sigma)}(u_i(z))$

Strategic Form

- The strategic form representation of the extensive form game is the normal form game defined by (N, S, u)
- A mixed strategy profile is a Nash equilibrium of the extensive form game if it constitutes a Nash equilibrium of its strategic form.

Behavior Strategies

- $b_i(h) \in \Delta(A(h))$: behavior strategy for player $i(h)$ at information set h
- $b_i(a \mid h)$: probability of action a at information set h
- behavior strategy $b_i \in \prod_{h \in H_i} \Delta(A(h))$
- independent mixing at each information set
- b_i outcome equivalent to the mixed strategy

$$\sigma_i(s_i) = \prod_{h \in H_i} b_i(s_i(h) \mid h)$$

- Is every mixed strategy equivalent to a behavior strategy?
- Yes, under perfect recall.

Perfect Recall

No player forgets any information he once had or actions he previously chose.

- If $x'' \in h(x')$, $x > x'$, and the same player i moves at both x and x' (and thus at x''), then there exists $\hat{x} \in h(x)$ (possibly $\hat{x} = x$) s.t. $\hat{x} > x''$ and the action taken at x along the path to x' is the same as the action taken at $\hat{x}$ along the path to x'' .
- x' and x'' distinguished by information i does not have, so he cannot have had it at $h(x)$
- x' and x'' consistent with the same action at $h(x)$ since i must remember his action there
- Equivalently, every node in $h \in H_i$ must be reached via the same sequence of i 's actions.

Equivalent Behavior Strategies

- $R_i(h) = \{s_i \mid h \text{ is on the path of } (s_i, s_{-i}) \text{ for some } s_{-i}\}$: set of i 's pure strategies that do not preclude reaching information set $h \in H_i$
- Under perfect recall, a mixed strategy σ_i is equivalent to a behavior strategy b_i defined by

$$b_i(a \mid h) = \frac{\sum_{\{s_i \in R_i(h) \mid s_i(h) = a\}} \sigma_i(s_i)}{\sum_{s_i \in R_i(h)} \sigma_i(s_i)}$$

when the denominator is positive.

Theorem 1 (Kuhn 1953)

In extensive form games with perfect recall, mixed and behavior strategies are outcome equivalent under the formulae (1) & (2).

Proof

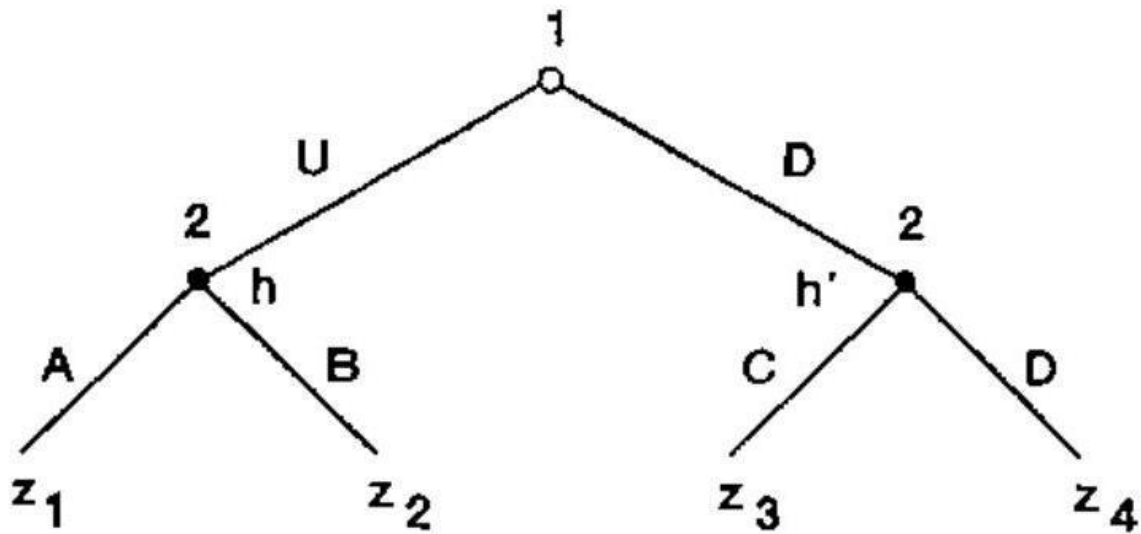
- $h_1, \dots, h_{\bar{k}}$: player i 's information sets preceding h in the tree
- Under perfect recall, reaching any node in h requires i to take the same action a_k at each h_k ,

$$R_i(h) = \{s_i \mid s_i(h_k) = a_k, \forall k = \overline{1, \bar{k}}\}.$$

- Conditional on getting to h , the distribution of continuation play at h is given by the relative probabilities of the actions available at h under the restriction of σ_i to $R_i(h)$,

$$b_i(a \mid h) = \frac{\{s_i \mid s_i(h_k) = a_k, \forall k = \overline{1, \bar{k}} \& s_i(h) = a\}}{\sum_{\{s_i \mid s_i(h_k) = a_k, \forall k = \overline{1, \bar{k}}\}} \sigma_i(s_i)}.$$

Example



Courtesy of The MIT Press. Used with permission.

Figure: Different mixed strategies can generate the same behavior strategy.

- $S_2 = \{(A, C), (A, D), (B, C), (B, D)\}$
- Consider the following mixed strategies:

$$\sigma_2 = 1/4(A, C) + 1/4(A, D) + 1/4(B, C) + 1/4(B, D)$$

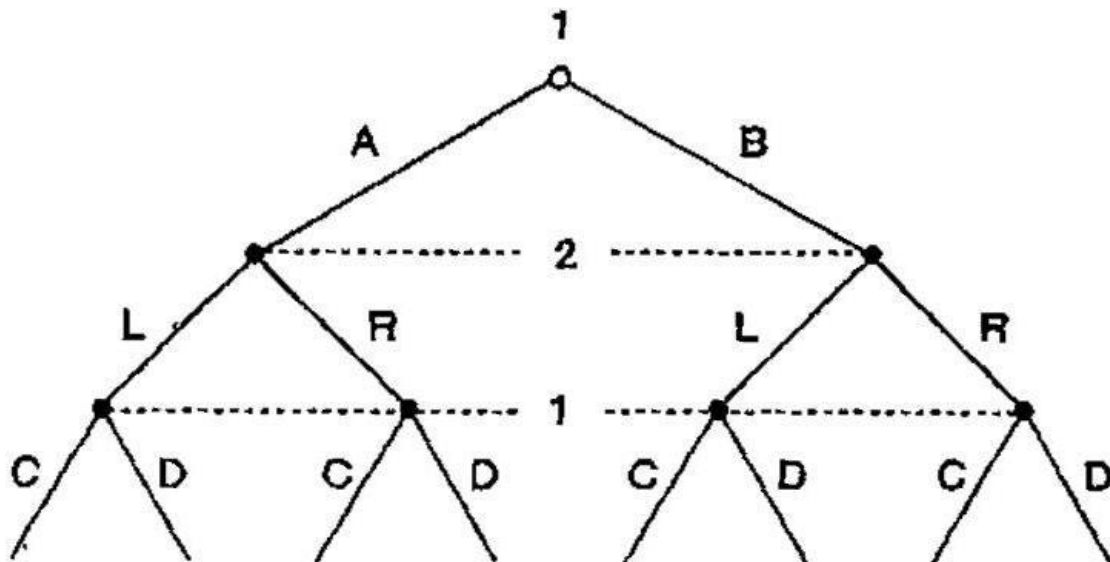
$$\sigma_2 = 1/2(A, C) + 1/2(B, D)$$

They generate - and are equivalent to - the behavior strategy b_2 with:

$$b_2(A \mid h) = b_2(B \mid h) = 1/2$$

$$b_2(C \mid h') = b_2(D \mid h') = 1/2$$

Example with Imperfect Recall

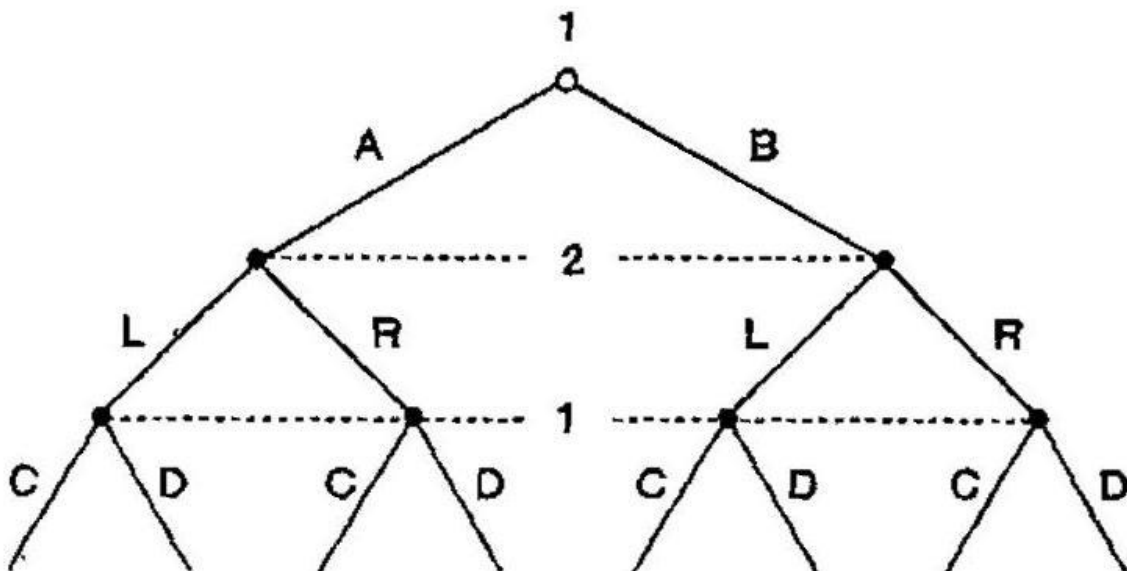


Courtesy of The MIT Press. Used with permission.

Figure: Player 1 forgets what he did at the initial node.

- $S_1 = \{(A, C), (A, D), (B, C), (B, D)\}$
- $\sigma_1 = 1/2(A, C) + 1/2(B, D) \rightarrow b_1 = (1/2A + 1/2B, 1/2C + 1/2D)$
- b_1 not equivalent to σ_1
- (σ_1, L) : prob. 1/2 for paths (A, L, C) and (B, L, D)
- (b_1, L) : prob. 1/4 to paths $(A, L, C), (A, L, D), (B, L, C), (B, L, D)$

Imperfect Recall and Correlations



Courtesy of The MIT Press. Used with permission.

- Since both A vs. B and C vs. D are choices made by player 1, the strategy σ_1 under which player 1 makes all his decisions at once allows choices at different information sets to be correlated
- Behavior strategies cannot produce this correlation, because when it comes time to choose between C and D , player 1 has forgotten whether he chose A or B .

Conventions

- Restrict attention to games with perfect recall, so we can use mixed and behavior strategies interchangeably.
- Behavior strategies are more convenient.
- Drop notation b for behavior strategies and denote by $\sigma_i(a \mid h)$ the probability with which player i chooses action a at information set h .

Backward Induction

- An extensive form game has perfect information if all information sets are singletons.
- Can solve games with perfect information using backward induction.
- Finite game $\rightarrow \exists$ penultimate nodes (successors are terminal nodes).
- The player moving at each penultimate node chooses an action that maximizes his payoff.
- Players at nodes whose successors are penultimate/terminal choose an optimal action given play at penultimate nodes.
- Work backwards to initial node...

Theorem 2 (Zermelo 1913; Kuhn 1953)

In a finite extensive form game of perfect information, the outcome(s) of backward induction constitutes a pure-strategy Nash equilibrium.

Stackelberg Duopoly (or Market Entrance)

- Incumbent firm 1 chooses output q (which is then fixed).
- A potential entrant, firm 2, observes q_1 and then chooses its output q_2 .
- Alternative interpretation: they choose capital levels instead of output.

- The profit for firm $i = 1, 2$ is:

$$\pi_i = q_i \cdot (1 - q_1 - q_2)$$

- Firm i produces output q_i , we use demand function.
- Each firm dislikes output by the other.
- A firm's marginal value of output decreases with the other's.
- Output choices are strategic substitutes.

Stackelberg Competition

- Profit maximization by firm 2 requires

$$q_2 = \frac{1 - q_1}{2}$$

- Firm 1 anticipates that firm 2 will act optimally, and therefore solves

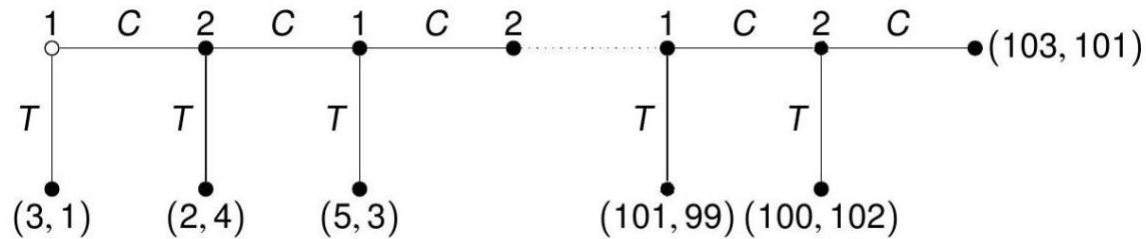
$$\max_{q_1} \left\{ q_1 \left(1 - q_1 - \frac{1 - q_1}{2} \right) \right\}$$

- Solution involves $q_1 = 1/2$, $q_2 = 1/4$, $\pi_1 = 1/8$, and $\pi_2 = 1/16$.
- Firm 1 has first mover advantage.
- In contrast, in the simultaneous move game, $q_1 = 1/3$, $q_2 = 1/3$, $\pi_1 = 1/9$, and $\pi_2 = 1/9$.

Centipede Game

- Player 1 has two piles in front of her: one contains 3 coins, the other 1.
- Player 1 can either take the larger pile and give the smaller one to player 2 (T) or push both piles across the table to player 2 (C).
- Every time the piles pass across the table, one coin is added to each.
- Players alternate in choosing whether to take the larger pile (T) or trust opponent with bigger piles (C).
- The game lasts 100 rounds.

What's the backward induction solution?



Chess Players and Backward Induction

- Palacios-Huerta and Volij (2009)
- Chess players and college students behave differently in the centipede game.
- Higher-ranked chess players end the game earlier.
- All Grandmasters in the experiment stopped at the first opportunity.
- Chess players are familiar with backward induction reasoning and need less learning to reach the equilibrium.
- Playing against non-chess-players, even chess players continue in the game longer.
- In long games, common knowledge of the ability to do complicated inductive reasoning becomes important for the prediction.

Subgame Perfection

- Backward induction solution is more than a Nash equilibrium.
- Actions are optimal given others' play-and form an equilibrium **starting at any intermediate node**: subgame perfection.
- It rules out non-credible threats.
- Subgame perfection extends backward induction to imperfect information games.
- Replace "smallest" subgames with a Nash equilibrium and iterate on the reduced tree (if there are multiple Nash equilibria in a subgame, all players expect same play).

Subgames

Subgame: part of a game that can be analyzed separately; strategically and informationally independent... information sets not "chopped up."

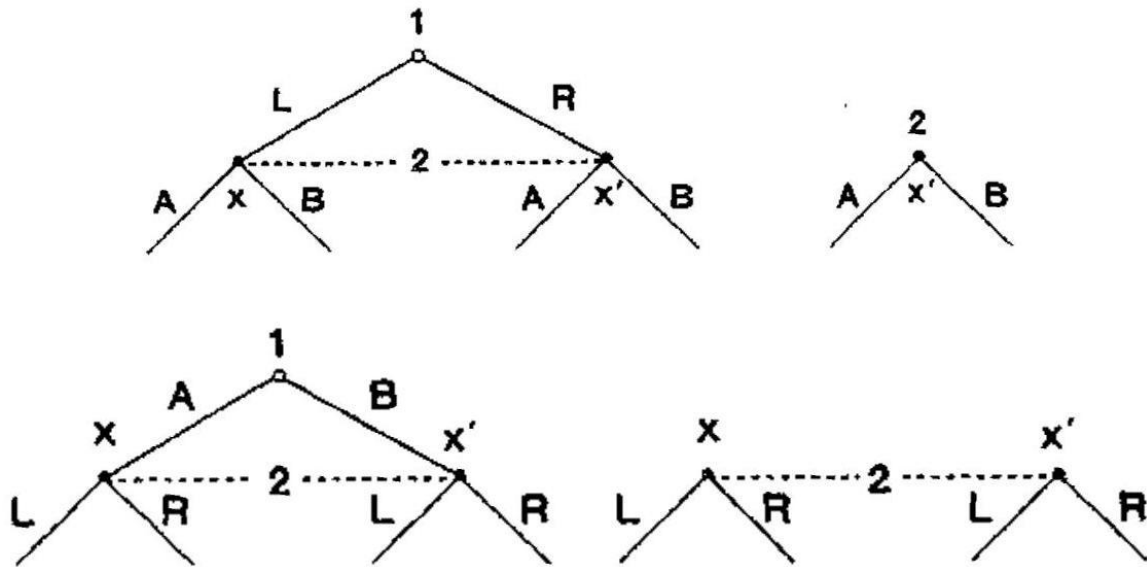
Definition 1

A subgame G of an extensive form game T is a subset of nodes such that:

- 1- The first node is unique (that is, it's not in a non-unitary information set)
- 2- It includes all the successors of the first node
- 3- If $x' \in G$ and $x'' \in h(x')$ then $x'' \in G$ (information sets are not chopped up)

The information sets, actions and payoffs in the subgame are inherited from T .

False Subgames



Courtesy of The MIT Press. Used with permission.

Subgame Perfect Equilibrium

σ : behavior strategy in T

- $\sigma \upharpoonright G$: the strategy profile induced by σ in subgame G of T (start play at the initial node of G , follow actions specified by σ , obtain payoffs from T at terminal nodes)
- Is $\sigma \upharpoonright G$ a Nash equilibrium of G for any subgame G ?

Definition 2

A strategy profile σ in an extensive form game T is a subgame perfect equilibrium if $\sigma \upharpoonright G$ is a Nash equilibrium of G **for every subgame G of T** .

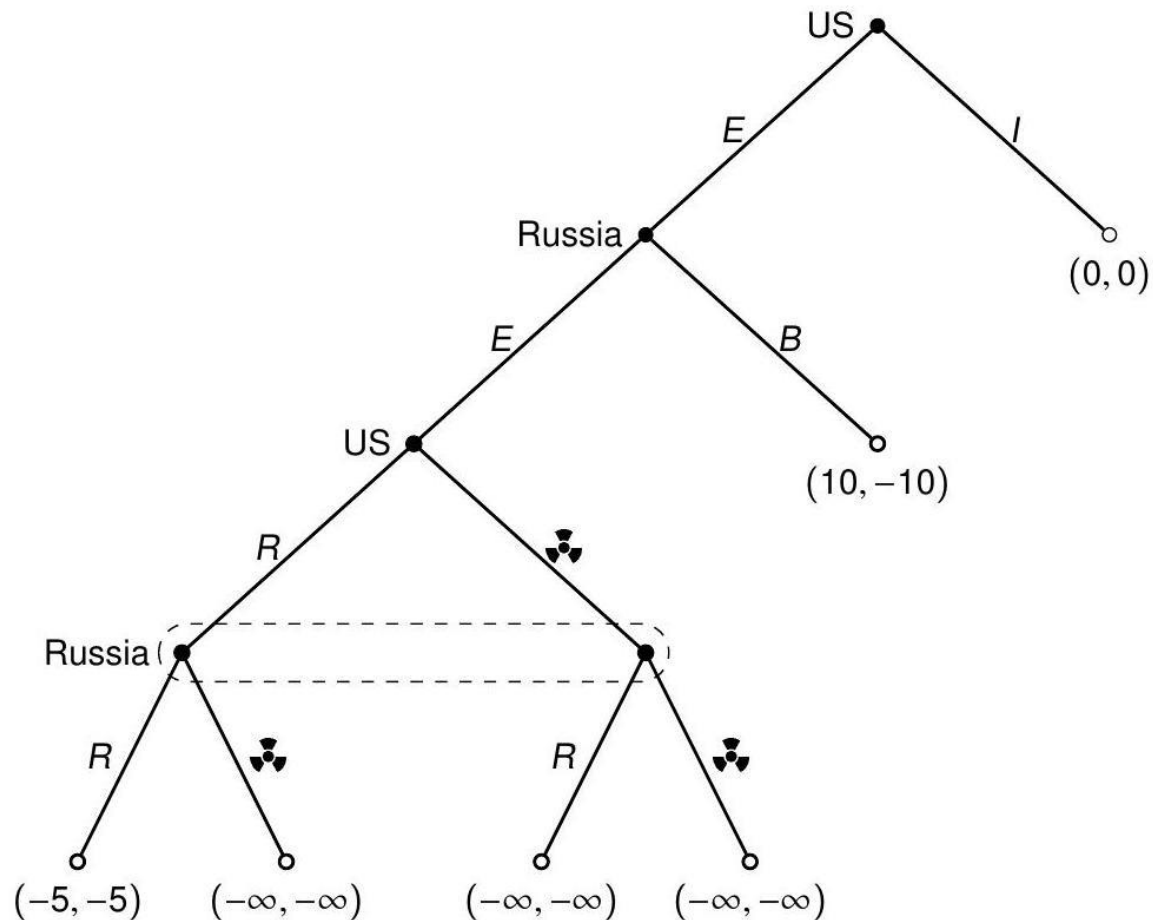
- Any game is a subgame of itself $\rightarrow$ a subgame perfect equilibrium is a Nash equilibrium.

- Subgame perfection coincides with backward induction in games of perfect information.

Nuclear Crisis

- Russia provokes the US...
- The U.S. can choose to escalate (E) or end the game by ignoring the provocation (I).
- If the game escalates, Russia faces a similar choice: to back down (B), but lose face, or escalate (E).
- Escalation leads to nuclear crisis: a simultaneous move game where each nation chooses to either retreat and lose credibility or detonate.
- Unless both countries retreat, retaliation to the first nuclear strike culminates in nuclear disaster, which is infinitely costly.

The Extensive Form

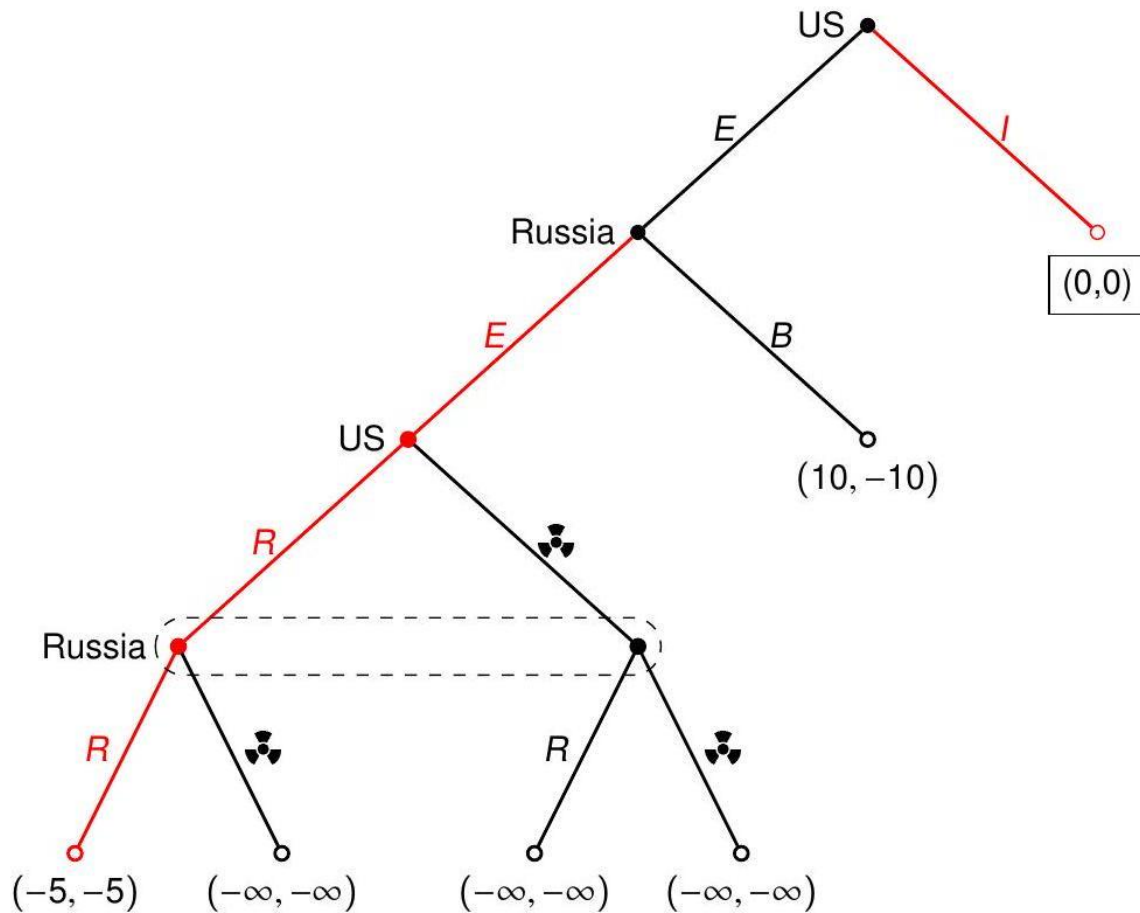


Last Stage

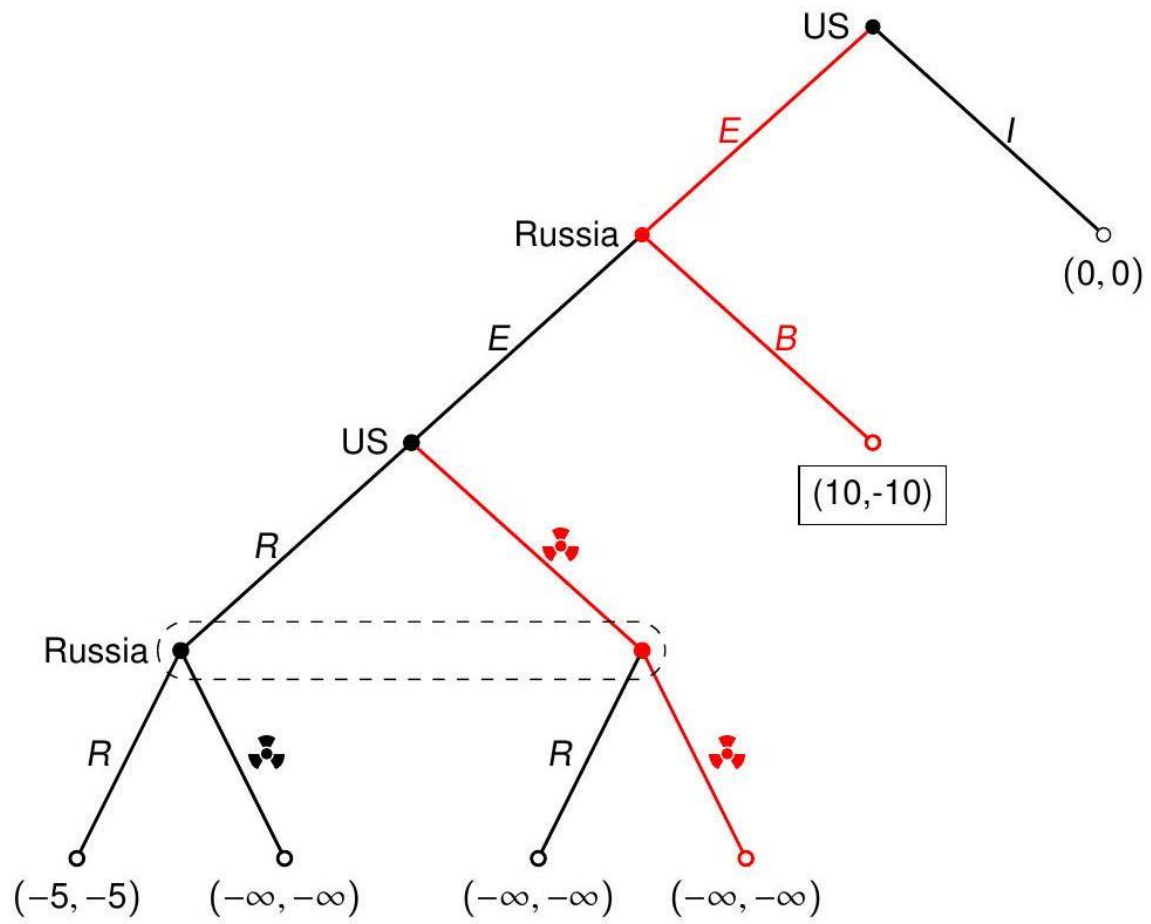
The simultaneous-move game at the last stage has two Nash equilibria.

	R	
R	$-5, 5^*$	$-\infty, -\infty$
	$-\infty, -\infty$	$-\infty, -\infty^*$

One Subgame Perfect Equilibrium



Another Subgame Perfect Equilibrium



Single-Deviation Principle and Bargaining

Multi-stage games with observable actions

- Finite set of players N
- Stages $t = 0, 1, 2, \dots$
- H : set of terminal histories (sequences of action profiles of possibly different lengths)
- At stage t , after having observed a non-terminal history of play $h_t = (a^0, \dots, a^{t-1}) \notin H$, each player i simultaneously chooses an action $a_i^t \in A_i(h_t)$
- $u_i(h)$: payoff of $i \in N$ for terminal history $h \in H$
- σ_i : behavior strategy for $i \in N$ specifies $\sigma_i(h) \in \Delta(A_i(h))$ for $h \notin H$

Often natural to identify "stages" with time periods.

Examples:

- Repeated games
- Alternating bargaining game

Unimprovable Strategies

To verify that a strategy profile σ constitutes a subgame perfect equilibrium (SPE) in a multi-stage game with observed actions...

It suffices to check whether there are any histories h_t where some player i can gain by deviating from playing $\sigma_i(h_t)$ at t and conforming to σ_i elsewhere.

$u_i(\sigma \mid h_t)$: expected payoff of player i in the subgame starting at h_t and played according to σ thereafter

Definition 1

A strategy σ_i is **unimprovable given σ_{-i}** if $u_i(\sigma_i, \sigma_{-i} \mid h_t) \geq u_i(\sigma'_i, \sigma_{-i} \mid h_t)$ for every h_t and σ'_i with $\sigma'_i(h) = \sigma_i(h)$ for all $h \neq h_t$.

Continuity at Infinity

If σ is an SPE then σ_i is unimprovable given σ_{-i} .

For the converse, we'll need an additional assumption:

Definition 2

A game is continuous at infinity if

$$\lim_{t \rightarrow \infty} \sup_{\{(h, \tilde{h}) | h_t = \tilde{h}_t\}} |u_i(h) - u_i(\tilde{h})| = 0, \forall i \in N$$

Events in the distant future are relatively unimportant.

Single (or One-Shot) Deviation Principle

Theorem 1

Consider a multi-stage game with observed actions **that is continuous at infinity**. If σ_i is unimprovable given σ_{-i} for all $i \in N$, then σ constitutes an SPE.

Proof allows for infinite action spaces at some stages. There exist versions for games with unobserved actions.

Proof

Suppose that σ_i is unimprovable given σ_{-i} , but σ_i is not a best response to σ_{-i} following some history h_t . Let σ_i^1 be a strictly better response and define

$$\varepsilon = u_i(\sigma_i^1, \sigma_{-i} | h_t) - u_i(\sigma_i, \sigma_{-i} | h_t) > 0$$

Since the game is continuous at infinity, there exists $t' > t$ and σ_i^2 s.t. σ_i^2 is identical to σ_i^1 at all information sets up to (and including) stage t' , σ_i^2 coincides with σ_i across all longer histories and

$$|u_i(\sigma_i^2, \sigma_{-i} | h_t) - u_i(\sigma_i^1, \sigma_{-i} | h_t)| < \varepsilon/2$$

Then

$$u_i(\sigma_i^2, \sigma_{-i} | h_t) > u_i(\sigma_i, \sigma_{-i} | h_t)$$

σ_i^3 : strategy obtained from σ_i^2 by replacing the stage t' actions following any history $h_{t'}$ with the corresponding actions under σ_i

Conditional on any $h_{t'}$, σ_i and σ_i^3 coincide, hence

$$u_i(\sigma_i^3, \sigma_{-i} \mid h_{t'}) = u_i(\sigma_i, \sigma_{-i} \mid h_{t'})$$

As σ_i is unimprovable given σ_{-i} , and conditional on $h_{t'}$ the subsequent play in strategies σ_i and σ_i^2 differs only at stage t' ,

$$u_i(\sigma_i, \sigma_{-i} \mid h_{t'}) \geq u_i(\sigma_i^2, \sigma_{-i} \mid h_{t'})$$

Then

$$u_i(\sigma_i^3, \sigma_{-i} \mid h_{t'}) \geq u_i(\sigma_i^2, \sigma_{-i} \mid h_{t'})$$

for all histories $h_{t'}$. Since σ_i^2 and σ_i^3 coincide before reaching stage t' ,

$$u_i(\sigma_i^3, \sigma_{-i} \mid h_t) \geq u_i(\sigma_i^2, \sigma_{-i} \mid h_t)$$

σ_i^4 : strategy obtained from σ_i^3 by replacing the stage $t' - 1$ actions following any history $h_{t'-1}$ with the corresponding actions under σ_i

Similarly,

$$u_i(\sigma_i^4, \sigma_{-i} \mid h_t) \geq u_i(\sigma_i^3, \sigma_{-i} \mid h_t) \dots$$

The final strategy $\sigma_i^{t'-t+3}$ is identical to σ_i conditional on h_t and

$$\begin{aligned} u_i(\sigma_i, \sigma_{-i} \mid h_t) &= u_i(\sigma_i^{t'-t+3}, \sigma_{-i} \mid h_t) \geq \dots \\ &\geq u_i(\sigma_i^3, \sigma_{-i} \mid h_t) \geq u_i(\sigma_i^2, \sigma_{-i} \mid h_t) > u_i(\sigma_i, \sigma_{-i} \mid h_t) \end{aligned}$$

a contradiction. QED.

FINITE AND INFINITE HORIZON BILATERAL BARGAINING

Finite Horizon Bilateral Bargaining

Two players, 1 and 2, bargain to determine the split of v dollars.

Rules:

Game begins in period 1: player 1 makes an offer of a split (a real number between 0 and v) to player 2, which player 2 may then accept or reject.

If she accepts, the proposed split is immediately implemented and the game ends.

If she rejects, nothing happens until period 2.

In period 2, the players' roles are reversed, with player 2 making an offer to player 1 and player 1 then being able to accept or reject it.

Each player has a discount factor of $\delta \in (0,1)$

A dollar received in period t is worth δ^{t-1} in period 1 dollars.

After some finite number of periods T , if an agreement has not yet been reached, the bargaining is terminated and the players each receive nothing.

Could be any exogenous payoff.

Unique subgame perfect Nash equilibrium (SPNE):

Suppose first that T is odd

So that player 1 makes the offer in period T if no previous agreement has been reached.

Player 2 is willing to accept any offer in this period because she will get zero if she refuses and the game is terminated

(She is indifferent about accepting an offer of zero).

Hence the unique SPNE in the subgame that begins in the final period when no agreement has been previously reached has player 1 offer player 2 zero and player 2 accept.

Payoffs from equilibrium play in this subgame are $(\delta^{T-1}v, 0)$.

Now consider play in the subgame starting in period $T - 1$ when no previous agreement has been reached.

Player 2 makes the offer in this period.

In any SPNE, player 1 will accept an offer in period $T - 1$ if and only if it provides her with a payoff of at least $\delta^{T-1}v$

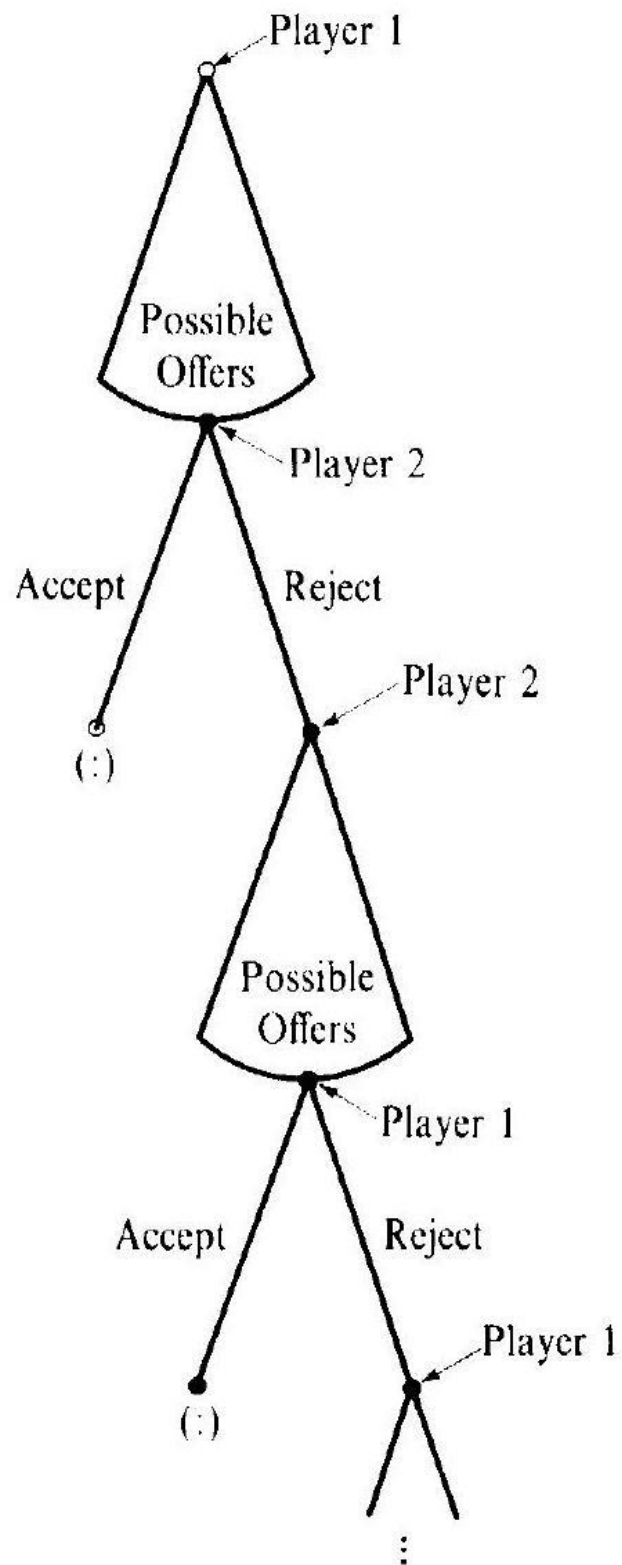
Otherwise she will do better rejecting it and waiting to make an offer in period T (she earns $\delta^{T-1}v$ by doing so).

Given this fact, in any SPNE, player 2 must make an offer in period $T - 1$ that gives player 1 a payoff of exactly $\delta^{T-1}v$

Player 1 accepts this offer

(Note that this is player 2's best offer among all those that would be accepted, and making an offer that will be rejected is worse for player 2 because it results in her receiving a payoff of zero).

The payoffs arising if the game reaches period $T - 1$ must therefore be $(\delta^{T-1}v, \delta^{T-2}v - \delta^{T-1}v)$.



Continuing in this fashion...

Unique SPNE when T is odd results in an agreement being reached in period 1 :

Payoff for player 1:

$$\begin{aligned} v_1^*(T) &= v[1 - \delta + \delta^2 - \dots + \delta^{T-1}] \\ &= v \left[(1 - \delta) \left(\frac{1 - \delta^{T-1}}{1 - \delta^2} \right) + \delta^{T-1} \right] \end{aligned}$$

Payoff to player 2:

$$v_2^*(T) = v - v_1^*(T).$$

If T is instead even:

Then player 1 must earn $v - \delta v_1^*(T - 1)$

... because in any SPNE, player 2 (who will be the first offerer in the odd-number-of-periods subgame that begins in period 2 if she rejects player 1's period 1 offer) will accept an offer in period 1 if and only if it gives her at least $\delta v_1^*(T - 1)$

... and player 1 will offer her exactly this amount.

As the number of periods grows large ($T \rightarrow \infty$)...

Player 1's payoff converges to $v/(1 + \delta)$

Player 2's payoff converges to $\delta v/(1 + \delta)$.

Infinite Horizon Bilateral Bargaining

We can no longer solve for the SPNE in this simple manner when the game has an infinite horizon.

In many games, introduction of an infinite horizon allows a broad range of behavior to emerge as subgame perfect.

Nevertheless, in the infinite horizon bargaining model, the SPNE concept is quite powerful.

There is a unique SPNE in this game

It is exactly the limiting outcome of the finite horizon model as $T \rightarrow \infty$

This model is due to Rubinstein (1982).

Claim: this game has a unique SPNE, with payoffs $v/(1 + \delta)$ and $\delta v/(1 + \delta)$.

Shaked and Sutton (1984): use the stationarity of the game

The subgame starting in period 2 looks exactly like that in period 1, but with the players' roles reversed.

To start, let $\bar{v}_1$ denote the largest payoff that player 1 gets in any SPNE

There may, in principle, be multiple SPNEs in this model

Given the stationarity of the model...

This is also the largest amount that player 2 can expect in the subgame that begins in period 2 after her rejection of player 1's period 1 offer...

...a subgame in which player 2 has the role of being the first player to make an offer.

Hence player 1's payoff in any SPNE cannot be lower than if player 2 gets the highest payoff in the next period

That is, player 1's payoff cannot be lower than $\underline{v}_1 = v - \delta \bar{v}_1$

If it was, then player 1 could do better by making a period 1 offer that gives player 2 just slightly more than $\delta \bar{v}_1$.

Player 2 is certain to accept any such offer because she will earn only $\delta \bar{v}_1$ by rejecting it

(note that we are using subgame perfection here, because we are requiring that the continuation of play after rejection is an SPNE in the continuation subgame and that player 2's response will be optimal given this fact).

Next claim: in any SPNE, $\bar{v}_1$ cannot be larger than $v - \delta \underline{v}_1$

Similar reasoning: player 2 won't accept any offer that gives him less than $\underline{v}_1$ tomorrow

Hence in any SPNE, player 2 is certain to reject any offer in period 1 that gives her less than $\delta \underline{v}_1$

... because she can earn at least $\delta \underline{v}_1$ by rejecting it and waiting to make an offer in period 2 .

Thus, player 1 can do no better than $v - \delta \underline{v}_1$ by making an offer that is accepted in period 1.

What about by making an offer that is rejected in period 1?

Since player 2 must earn at least $\delta \underline{v}_1$ if this happens, and since agreement cannot occur before period 2, player 1 can earn no more than $\delta v - \delta \underline{v}_1$ by doing this.

In short, we have $\bar{v}_1 \leq v - \delta \underline{v}_1$

Remember that we defined:

$$\underline{v}_1 = v - \delta \bar{v}_1$$

Rearrange it to get:

$$v = \underline{v}_1 + \delta \bar{v}_1$$

Plug into previous inequality:

$$\bar{v}_1 \leq v - \delta \underline{v}_1 = (\underline{v}_1 + \delta \bar{v}_1) - \delta \underline{v}_1$$

$$\bar{v}_1 - \delta \bar{v}_1 \leq \underline{v}_1 - \delta \underline{v}_1$$

$$\bar{v}_1(1 - \delta) \leq \underline{v}_1(1 - \delta)$$

$$\bar{v}_1 \leq \underline{v}_1$$

Given the definitions of $\underline{v}_1$ and $\bar{v}_1$, we must also have:

$$\bar{v}_1 \geq \underline{v}_1$$

Hence:

$$\underline{v}_1 = \bar{v}_1$$

So player 1 's SPNE payoff is uniquely determined.

Denote this payoff by v_1° .

We again the fact that $\underline{v}_1 = v - \delta \bar{v}_1$:

$$v_1^\circ = v - \delta v_1^\circ$$

Hence:

$$v_1^\circ = v/(1 + \delta)$$

Player 2 must earn:

$$v_2^\circ = v - v_1^\circ = \delta v/(1 + \delta)$$

An agreement will be reached in the first period:

Player 1 will find it worthwhile to make an offer that player 2 accepts.

SPNE strategies:

A player who has just received an offer accepts it if and only if she is offered at least δv_1^o

A player whose turn it is to make an offer offers exactly δv_1^o to the player receiving the offer.

Remarks:

The coincidence of the infinite horizon equilibrium with the limit of the finite horizon equilibria in this model is not a general property of infinite horizon games.

The outcomes of game-theoretic models of bargaining can be quite sensitive to the precise specification of the bargaining process and players' preferences.

Bargaining with Alternating Offers: full analysis

Rubinstein (1982)

- players $i = 1, 2; j = 3 - i$
- set of feasible utility pairs

$$U = \{(u_1, u_2) \in [0, \infty)^2 \mid u_2 \leq g_2(u_1)\}$$

- g_2 s. decreasing, concave (and hence continuous), $g_2(0) > 0$
- δ_i : discount factor of player i
- at every time $t = 0, 1, \dots$, player $i(t)$ proposes an alternative $u = (u_1, u_2) \in U$ to player $j(t) = 3 - i(t)$

$$i(t) = \begin{cases} 1 & \text{for } t \text{ even} \\ 2 & \text{for } t \text{ odd} \end{cases}$$

- if $j(t)$ accepts the offer, game ends yielding payoffs $(\delta_1^t u_1, \delta_2^t u_2)$
- otherwise, game proceeds to period $t + 1$

Stationary SPE

Define $g_1 = g_2^{-1}$.

Graphs of g_2 and g_1^{-1} : Pareto-frontier of U

Let (m_1, m_2) be the unique solution to the following system of equations

$$\begin{aligned} m_1 &= \delta_1 g_1(m_2) \\ m_2 &= \delta_2 g_2(m_1) \end{aligned}$$

(m_1, m_2) is the intersection of the graphs of $\delta_2 g_2$ and $(\delta_1 g_1)^{-1}$.

SPE in "stationary" strategies: in any period where player i has to make an offer to j , he offers u with $u_j = m_j$ and $u_i = g_i(m_j)$, and j accepts only offers u with $u_j \geq m_j$.

Single-deviation principle: constructed strategies form an SPE.

Is the SPE unique?

Iterated Conditional Dominance

Definition 3

In a multi-stage game with observable actions, an action a_i is **conditionally dominated** at stage t given history h_t if, in the subgame starting at h_t , every strategy for player i that assigns positive probability to a_i is strictly dominated.

Proposition 1

In any multi-stage game with observable actions, every SPE survives the iterated elimination of conditionally dominated strategies.

Equilibrium uniqueness

Iterated conditional dominance: stationary equilibrium is essentially the unique SPE.

Theorem 2

The SPE of the alternating-offer bargaining game is unique, except for the decision to accept or reject Pareto-inefficient offers.

Proof

- Following a disagreement at date t , player i cannot obtain a period t expected payoff greater than

$$M_i^0 = \delta_i \max_{u \in U} u_i = \delta_i g_i(0)$$

- Rejecting an offer u with $u_i > M_i^0$ is conditionally dominated by accepting such an offer for i .
- Once we eliminate dominated actions, i accepts all offers u with $u_i > M_i^0$ from j .
- Making any offer u with $u_i > M_i^0$ is dominated for j by an offer $\bar{u} = \lambda u + (1 - \lambda)(M_i^0, g_j(M_i^0))$ for $\lambda \in (0,1)$ (both offers are accepted immediately).

Under the surviving strategies

- j can reject an offer from i and make a counteroffer next period that leaves him with slightly less than $g_j(M_i^0)$, which i accepts; it is conditionally dominated for j to accept any offer smaller than

$$m_j^1 = \delta_j g_j(M_i^0)$$

- i cannot expect to receive a continuation payoff greater than

$$M_i^1 = \max(\delta_i g_i(m_j^1), \delta_i^2 M_i^0) = \delta_i g_i(m_j^1)$$

after rejecting an offer from j

$$\delta_i g_i(m_j^1) = \delta_i g_i(\delta_j g_j(M_i^0)) \geq \delta_i g_i(g_j(M_i^0)) = \delta_i M_i^0 \geq \delta_i^2 M_i^0$$

Recursively define

$$\begin{aligned} m_j^{k+1} &= \delta_j g_j(M_i^k) \\ M_i^{k+1} &= \delta_i g_i(m_j^{k+1}) \end{aligned}$$

for $i = 1, 2$ and $k \geq 1$. $(m_i^k)_{k \geq 0}$ is increasing and $(M_i^k)_{k \geq 0}$ is decreasing.

Prove by induction on k that, under any strategy that survives iterated conditional dominance, player $i = 1, 2$

- never accepts offers with $u_i < m_i^k$
- always accepts offers with $u_i > M_i^k$, but making such offers is dominated for j .
- The sequences (m_i^k) and (M_i^k) are monotonic and bounded, so they need to converge. The limits satisfy

$$\begin{aligned} m_j^\infty &= \delta_j g_j(\delta_i g_i(m_j^\infty)) \\ M_i^\infty &= \delta_i g_i(m_j^\infty) \end{aligned}$$

- (m_1^∞, m_2^∞) is the (unique) intersection point of the graphs of the functions $\delta_2 g_2$ and $(\delta_1 g_1)^{-1}$
- $M_i^\infty = \delta_i g_i(m_j^\infty) = m_i^\infty$
- All strategies of i that survive iterated conditional dominance accept u with $u_i > M_i^\infty = m_i^\infty$ and reject u with $u_i < m_i^\infty = M_i^\infty$.

In an SPE

- at any history where i is the proposer, i 's payoff is at least $g_i(m_j^\infty)$: offer u arbitrarily close to $(g_i(m_j^\infty), m_j^\infty)$, which j accepts under the strategies surviving the elimination process
- i cannot get more than $g_i(m_j^\infty)$
- any offer made by i specifying a payoff greater than $g_i(m_j^\infty)$ for himself would leave j with less than m_j^∞ ; such offers are rejected by j under the surviving strategies
- under the surviving strategies, j never offers i more than

$$M_i^\infty = \delta_i g_i(m_j^\infty) \leq g_i(m_j^\infty)$$

- hence i 's payoff at any history where i is the proposer is exactly $g_i(m_j^\infty)$; possible only if i offers $(g_i(m_j^\infty), m_j^\infty)$ and j accepts with probability 1
Uniquely pinned down actions at every history, except those where j has just received an offer (u_i, m_j^∞) for some $u_i < g_i(m_j^\infty)$...

QED

Properties of the equilibrium

- The SPE is efficient-agreement is obtained in the first period, without delay.
- SPE payoffs: $(g_1(m_2), m_2)$, where (m_1, m_2) solve

$$m_1 = \delta_1 g_1(m_2)$$

$$m_2 = \delta_2 g_2(m_1)$$
- Patient players get higher payoffs: the payoff of player i is increasing in δ_i and decreasing in δ_j .
- For a fixed $\delta_1 \in (0,1)$, the payoff of player 2 converges to 0 as $\delta_2 \rightarrow 0$ and to $\max_{u \in U} u_2$ as $\delta_2 \rightarrow 1$.
- If U is symmetric and $\delta_1 = \delta_2$, player 1 enjoys a first mover advantage: $m_1 = m_2$ and $g_1(m_2) = m_2/\delta > m_2$.

Bargaining with random selection of proposer

- Two players need to divide \$1.
- Every period $t = 0, 1, \dots$ player 1 is chosen with probability p to make an offer to player 2.
- Player 2 accepts or rejects 1's proposal.
- Roles are interchanged with probability $1 - p$.
- In case of disagreement the game proceeds to the next period.
- The game ends as soon as an offer is accepted.
- Player $i = 1, 2$ has discount factor δ_i .

Equilibrium

- The unique equilibrium is stationary, i.e., each player i has the same expected payoff v_i in every subgame.
- Payoffs solve

$$v_1 = p \cdot (1 - \delta_2 v_2) + (1 - p) \cdot \delta_1 v_1$$

$$v_2 = p \cdot \delta_2 v_2 + (1 - p) \cdot (1 - \delta_1 v_1)$$

- The solution is

$$v_1 = \frac{1}{1 + \frac{(1-p)(1-\delta_1)}{p(1-\delta_2)}}$$

$$v_2 = \frac{1}{1 + \frac{p(1-\delta_2)}{(1-p)(1-\delta_1)}}$$

- Immediate agreement
- First mover advantage
- v_1 increases with p , v_2 decreases with p .
- For $\delta_1 = \delta_2$, we obtain $v_1 = p, v_2 = 1 - p$.
- Patience pays off
- v_i increases with δ_i and decreases with δ_j .
- Fix δ_j and take $\delta_i \rightarrow 1$, we get $v_i \rightarrow 1$ and $v_j \rightarrow 0$.

Repeated Games

Consider a static game $\Gamma_N = [I, \{S_i\}_i, \{u_i\}_i]$.

In a repeated game, players play this game in periods $t = 1, 2, \dots, T$, with either $T < \infty$ or $T = \infty$.

We have the following results:

1. Playing the static NE in every period is a SPNE.
2. If Γ_N has only one Nash Equilibrium and $T < \infty$, then the only SPNE of the repeated game is the repetition of the static equilibrium in every period.

We interpret this as the impossibility of cooperation (ie, coordinating on something different from the NE).

Proof is simple and follow from the impossibility of credible rewards (for cooperation) or punishments (for non-cooperation).

Games with multiple Nash equilibria

If the static game has multiple equilibria, it may be possible to build different SPNE.

Example:

	A	B	C	D
A	3,3	1,1	1,1	1,1
B	1,1	7,7	1,8	1,1
C	1,1	8,1	1,1	1,1
D	1,1	1,1	1,1	5,5

Assume this game is played twice.

SPNE: play B in the first period; if (B, B) was the outcome of the first period, play D in the second period; otherwise play A .

Infinitely Repeated Games

Threats and promises about future behavior can influence current behavior in repeated relationships.

Reasoning a bit different from finite games

No final state to start backward induction at.

Finite-horizon:

If there are multiple Nash equilibria of the stage game G ...

then there may be subgame-perfect outcomes of the repeated game $G(T)$ in which, for any $t < T$, the outcome of stage t is not a Nash equilibrium of G .

Infinitely repeated games:

even if the stage game has a unique Nash equilibrium...

there may be subgame-perfect outcomes of the infinitely repeated game in which **no stage's outcome is a Nash equilibrium of G** .

There is always a future period where punishment can be enacted.

Begin by studying the infinitely repeated Prisoners' Dilemma.

Then consider following class of infinitely repeated games:

a static game of complete information, G , is repeated infinitely

the outcomes of all previous stages observed before the current stage begins

The Prisoners' Dilemma below is to be repeated infinitely

For each t , the outcomes of the $t - 1$ preceding plays of the stage game are observed before the t^{th} stage begins.

	L_2	R_2
L_1	1,1	5,0
R_1	0,5	4,4

Payoffs in infinitely-repeated games and the discount factor

Summing the payoffs from this infinite sequence of stage games does not provide a useful measure of a player's payoff in the infinitely repeated game.

Receiving a payoff of 4 in every period is better than receiving a payoff of 1 in every period

But the sum of the payoffs is infinity in both cases.

We will use a discount factor $\delta \in (0,1)$

For example, $\delta = 1/(1 + r)$ is the value today of a dollar to be received one stage later, where r is the interest rate per stage.

We can also use δ to reinterpret what we call an infinitely repeated game as a **repeated game that ends after a random number of repetitions**.

Suppose that after each stage is played a (weighted) coin is flipped to determine whether the game will end.

If the probability is p that the game ends immediately, then a payoff π to be received in the next stage (if it is played) is worth only $(1 - p)\pi/(1 + r)$ before this stage's coin flip occurs.

Let $\delta = (1 - p)/(1 + r)$.

Then the present value $\pi_1 + \delta\pi_2 + \delta^2\pi_3 + \dots$ reflects both the time-value of money and the possibility that the game will end.

Given a discount factor and a player's payoffs from an infinite sequence of stage games...

We can compute the present value of the payoffs: the lump-sum payoff that could be put in the bank now so as to yield the same bank balance at the end of the sequence.

Definition Given the discount factor δ , the present value of the infinite sequence of payoffs $\pi_1, \pi_2, \pi_3, \dots$ is:

$$\pi_1 + \delta\pi_2 + \delta^2\pi_3 + \dots = \sum_{t=1}^{\infty} \delta^{t-1}\pi_t$$

Example: SPNE in the repeated Prisoners' Dilemma

Consider the infinitely repeated Prisoners' Dilemma in which each player's discount factor is δ

Each player's payoff in the repeated game is the present value of the player's payoffs from the stage games.

We will show that cooperation (R_1, R_2) can occur in every stage of a subgame-perfect outcome of the infinitely repeated game, even though the only Nash equilibrium in the stage game is noncooperation—that is, (L_1, L_2) .

The argument is in the spirit of our analysis of the two-stage repeated game with multiple Nash Equilibria

If the players cooperate today then they play a high-payoff equilibrium tomorrow

Otherwise they play a low-payoff equilibrium tomorrow.

The difference is that the high-payoff equilibrium that might be played tomorrow is continued cooperation tomorrow and thereafter (instead of 'good NE from stage game').

Player i begins the infinitely repeated game by cooperating and then cooperates in each subsequent stage game if and only if both players have cooperated in every previous stage.

Formally, player i 's strategy is:

Play R_i in the first stage.

In the t^{th} stage, if the outcome of all $t - 1$ preceding stages has been (R_1, R_2) then play R_i

Otherwise, play L_i .

This strategy is an example of a **trigger strategy**, so called because player i cooperates until someone fails to cooperate, which triggers a switch to noncooperation forever after.

If both players adopt this trigger strategy then the outcome of the infinitely repeated game will be (R_1, R_2) in every stage.

If δ is close enough to one then it is a Nash equilibrium of the infinitely repeated game for both players to adopt this strategy.

Such a Nash equilibrium is subgame-perfect, in a sense to be made precise.

Proposition: the trigger strategy is a Nash Equilibrium (in the whole game).

To show that it is a Nash equilibrium of the infinitely repeated game for both players to adopt the trigger strategy...

Assume that player i has adopted the trigger strategy

Then show that, provided δ is close enough to one, it is a best response for player j to adopt the strategy also.

We will use the one-shot deviation property here.

That is, we'll only consider deviations in one period.

Since player i will play L_i forever once one stage's outcome differs from (R_1, R_2) , player j 's best response is indeed to play L_j forever once one stage's outcome differs from (R_1, R_2) .

It remains to determine player j 's best response in the first stage, and in any stage such that all the preceding outcomes have been (R_1, R_2) .

Payoff of deviation, or non-cooperation:

Playing L_j will yield a payoff of 5 this stage but will trigger noncooperation by player i (and therefore also by player j) forever after, so the payoff in every future stage will be 1.

The present value of this sequence of payoffs is

$$5 + \delta \cdot 1 + \delta^2 \cdot 1 + \dots = 5 + \frac{\delta}{1 - \delta}$$

Payoff of cooperation:

Playing R_j will yield a payoff of 4 in this stage and will lead to exactly the same choice between L_j and R_j in the next stage.

One way to compute the present value of this:

$$4 + \delta \cdot 4 + \delta^2 \cdot 4 + \dots = \frac{1}{1 - \delta} \cdot 4 = \frac{4}{1 - \delta}$$

Another way to compute it:

Let V denote the present value of the infinite sequence of payoffs player j receives from making this choice optimally (now and every time it arises subsequently).

If playing R_j is optimal, then:

$$V = 4 + \delta V$$

$$V = 4/(1 - \delta)$$

Because playing R_j leads to the same choice next stage.

Condition for cooperation in first stage:

Playing R_j is optimal if and only if

$$\frac{4}{1 - \delta} \geq 5 + \frac{\delta}{1 - \delta} \quad (2.3.1)$$

$$\delta \geq 1/4$$

In the first stage, and in any stage such that all the preceding outcomes have been (R_1, R_2) , player j 's optimal action (given that player i has adopted the trigger strategy) is R_j if and only if $\delta \geq 1/4$.

Combining this observation with the fact that j 's best response is to play L_j forever once one stage's outcome differs from (R_1, R_2) , we have that:

It is a Nash equilibrium for both players to play the trigger strategy if and only if $\delta \geq 1/4$.

That is: choosing R is optimal if and only if players are patient enough.

If so, the short-term benefit is offset by the long-term loss.

General case

Define now a strategy in a repeated game, a subgame in a repeated game, and a subgame-perfect Nash equilibrium in a repeated game.

Stage game $G = \{A_1, \dots, A_n; u_1, \dots, u_n\}$:

Static game of complete information in which players 1 through n simultaneously choose actions a_1 through a_n from the action spaces A_1 through A_n , respectively, and payoffs are $u_1(a_1, \dots, a_n)$ through $u_n(a_1, \dots, a_n)$.

We had before a finitely repeated game $G(T)$ based on G .

Define the analogous infinitely repeated game.

Definition: infinitely repeated game

$G(\infty, \delta)$ denotes the infinitely repeated game in which G is repeated forever and the players share the discount factor δ .

For each t , the outcomes of the $t - 1$ preceding plays of the stage game are observed before the t^{th} stage begins.

Each player's payoff in $G(\infty, \delta)$ is the present value of the player's payoffs from the infinite sequence of stage games.

In any game (repeated or otherwise), a player's strategy is a complete plan of action:

It specifies a feasible action for the player in every contingency in which the player might be called upon to act.

Remember that in a dynamic game, a strategy is more complicated than a simple action.

In the finitely repeated game $G(T)$ or the infinitely repeated game $G(\infty, \delta)$, the history of play through stage t is the record of the players' choices in stages 1 through t .

The players might have chosen $(a_{11}, \dots, a_{n1})$ in stage 1, $(a_{12}, \dots, a_{n2})$ in stage 2, ..., and $(a_{1t}, \dots, a_{nt})$ in stage t , for example, where for each player i and stage s the action a_{is} belongs to the action space A_i .

Definition In the finitely repeated game $G(T)$ or the infinitely repeated game $G(\infty, \delta)$, a player's strategy specifies the action the player will take in each stage, **for each possible history of play through the previous stage**.

Subgames

Definition In the finitely repeated game $G(T)$, a subgame beginning at stage $t + 1$ is the repeated game in which G is played $T - t$ times, denoted $G(T - t)$.

There are many subgames that begin at stage $t + 1$, **one for each of the possible histories of play through stage t** .

In the infinitely repeated game $G(\infty, \delta)$, each subgame beginning at stage $t + 1$ is identical to the original game $G(\infty, \delta)$.

As in the finite-horizon case, there are as many subgames beginning at stage $t + 1$ of $G(\infty, \delta)$ as there are possible histories of play through stage t .

Note that the t^{th} stage of a repeated game taken on its own is not a subgame of the repeated game (assuming $t < T$ in the finite case).

Proposition: The trigger strategy in the PD is a subgame perfect Nash equilibrium.

We must show that the trigger strategies constitute a Nash equilibrium on every subgame of that infinitely repeated game.

Recall that every subgame of an infinitely repeated game is identical to the game as a whole.

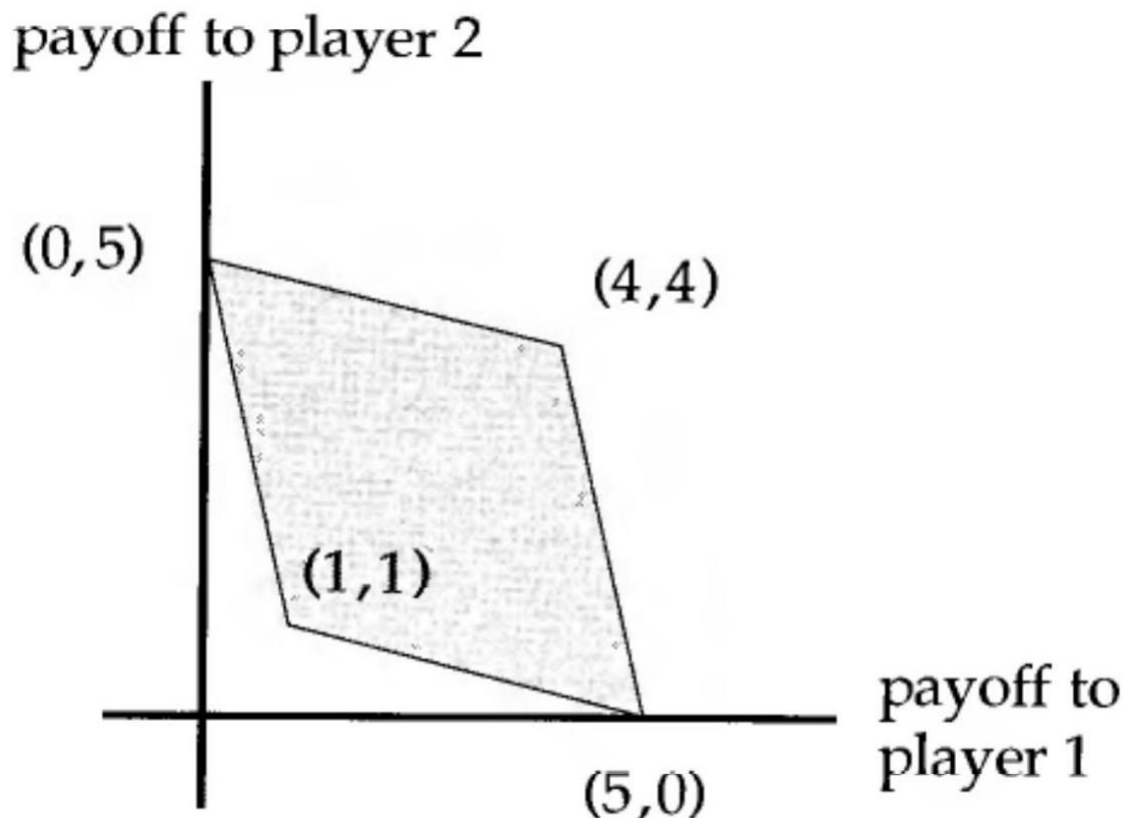
In the trigger-strategy Nash equilibrium of the infinitely repeated Prisoners' Dilemma, these subgames can be grouped into two classes:

- (i) Subgames in which all the outcomes of earlier stages have been (R_1, R_2)
- (ii) Subgames in which the outcome of at least one earlier stage differs from (R_1, R_2) .

If the players adopt the trigger strategy for the game as a whole, then

- (i) the players' strategies in a subgame in the first class are again the trigger strategy, which we have shown to be a Nash equilibrium of the game as a whole
- (ii) and the players' strategies in a subgame in the second class are simply to repeat the stage-game equilibrium (L_1, L_2) forever, which is also a Nash equilibrium of the game as a whole.

Thus, the trigger-strategy Nash equilibrium of the infinitely repeated Prisoners' Dilemma is subgame-perfect.



We next apply analogous arguments in the infinitely repeated game $G(\infty, \delta)$.

These arguments lead to Friedman's (1971) Theorem for infinitely repeated games.

To state the theorem, we need two final definitions.

First, we call the payoffs $(x_1, \dots, x_n)$ **feasible** in the stage game G if they are a convex combination (i.e., a weighted average, where the weights are all nonnegative and sum to one) of the pure-strategy payoffs of G .

The set of feasible payoffs for the Prisoners' Dilemma we're considering is the shaded region in the figure above.

The pure-strategy payoffs (1,1), (0,5), (4,4), and (5,0) are feasible.

Other feasible payoffs include the pairs (x, x) for $1 < x < 4$, which result from weighted averages of (1,1) and (4,4), and the pairs (y, z) for $y + z = 5$ and $0 < y < 5$, which result from weighted averages of (0,5) and (5,0).

The other pairs in (the interior of) the shaded region above are weighted averages of more than two pure-strategy payoffs.

To achieve a weighted average of pure-strategy payoffs, the players could use a public randomizing device: by playing (L_1, R_2) or (R_1, L_2) depending on a flip of a (fair) coin, for example, they achieve the expected payoffs (2.5, 2.5).

Second: we need a rescaling of the players' payoffs.

We continue to define each player's payoff in the infinitely repeated game $G(\infty, \delta)$ to be the present value of the player's infinite sequence of stagegame payoffs

But it is more convenient to express this present value in terms of the **average** payoff from the same infinite sequence of stage-game payoffs-the payoff that would have to be received in every stage so as to yield the same present value.

Let the discount factor be δ .

Suppose the infinite sequence of payoffs $\pi_1, \pi_2, \pi_3, \dots$ has a present value of V .

If the payoff π were received in every stage, the present value would be $\pi/(1 - \delta)$.

For π to be the average payoff from the infinite sequence $\pi_1, \pi_2, \pi_3 \dots$ with discount factor δ , these two present values must be equal, so $\pi = V(1 - \delta)$.

That is, **the average payoff is $(1 - \delta)$ times the present value.**

Definition Given the discount factor δ , the average payoff of the infinite sequence of payoffs $\pi_1, \pi_2, \pi_3, \dots$ is

$$(1 - \delta) \sum_{t=1}^{\infty} \delta^{t-1} \pi_t$$

The advantage of the average payoff over the present value is that **the former is directly comparable to the payoffs from the stage game.**

In the Prisoners' Dilemma we're considering, both players might receive a payoff of 4 in every period.

Such an infinite sequence of payoffs has an average payoff of 4 but a present value of $4/(1 - \delta)$.

Since the average payoff is just a rescaling of the present value, however, **maximizing the average payoff is equivalent to maximizing the present value.**

We are at last ready to state the main result in our discussion of infinitely repeated games:

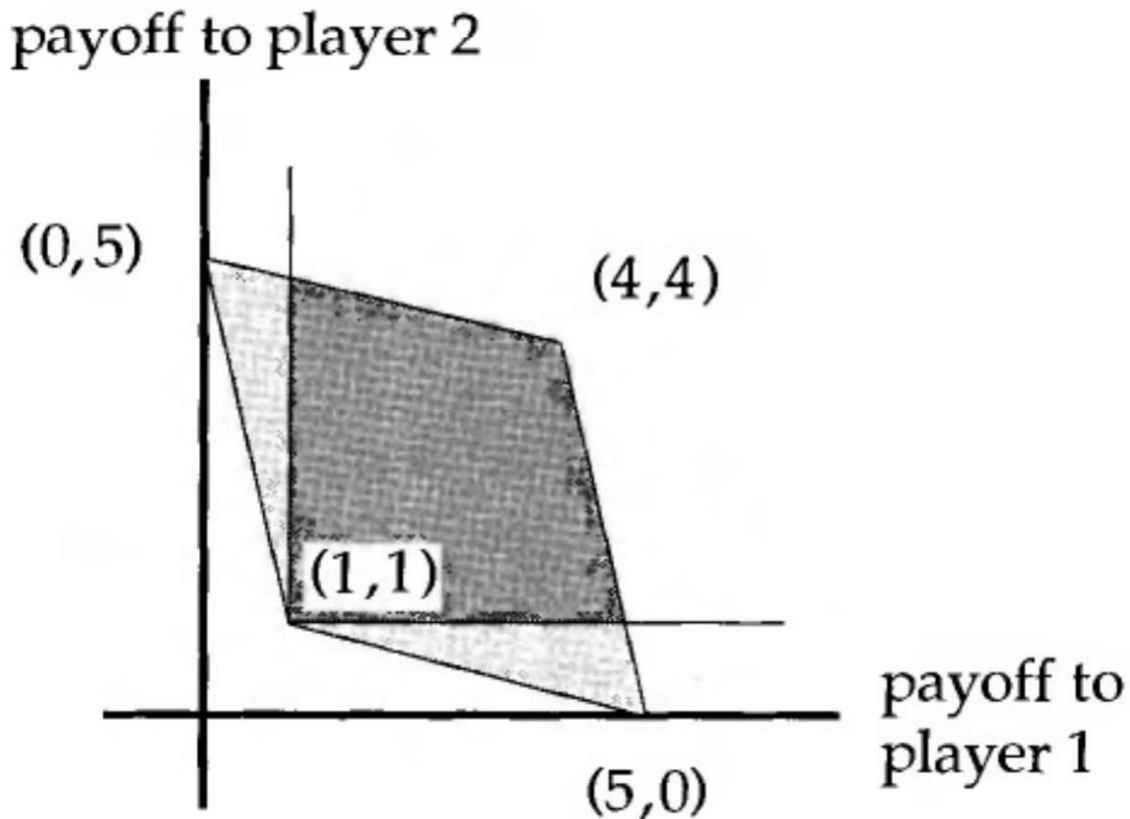
Theorem (Friedman 1971)

Let G be a finite, static game of complete information.

Let $(e_1, \dots, e_n)$ denote the payoffs from a Nash equilibrium of G

Let $(x_1, \dots, x_n)$ denote any other feasible payoffs from G .

If $x_i > e_i$ for every player i and if δ is sufficiently close to one, then there exists a subgame-perfect Nash equilibrium of the infinitely repeated game $G(\infty, \delta)$ that achieves $(x_1, \dots, x_n)$ as the average payoff.



In the context of our Prisoners' Dilemma ...

Friedman's Theorem guarantees that any point in the cross-hatched region in the figure above can be achieved...

as the average payoff in a subgame-perfect Nash equilibrium of the repeated game...

provided the discount factor is sufficiently close to one.

One remark

In the (one-shot) Prisoners' Dilemma, player i can guarantee receiving at least the Nash equilibrium payoff of 1 by playing L_i .

This is particular to PD.

In a one-shot Cournot duopoly game, in contrast, a firm cannot guarantee receiving the Nash-equilibrium profit by producing the Nash-equilibrium quantity

The only profit a firm can guarantee receiving is zero, by producing zero

Given an arbitrary stage game G , let r_i denote player i 's reservation payoff-the largest payoff player i can guarantee receiving, no matter what the other players do.

Hence $r_i \leq e_i$ (where e_i is player i 's Nash equilibrium payoff used in Friedman's Theorem)

If r_i were greater than e_i , it would not be a best response for player i to play his or her Nash-equilibrium strategy.

In the Prisoners' Dilemma, $r_i = e_i$, but in the Cournot duopoly game (and typically), $r_i < e_i$.

Fudenberg and Maskin (1986) show that for two-player games, the reservation payoffs (r_1, r_2) can replace the equilibrium payoffs (e_1, e_2) in the statement of Friedman's Theorem.

That is, if (x_1, x_2) is a feasible payoff from G , with $x_i > r_i$ for each i , then for δ sufficiently close to one there exists a subgame-perfect Nash equilibrium of $G(\infty, \delta)$ that achieves (x_1, x_2) as the average payoff, even if $x_i < e_i$ for one or both of the players.

For games with more than two players, Fudenberg and Maskin provide a mild condition under which the reservation payoffs $(r_1, \dots, r_n)$ can replace the equilibrium payoffs $(e_1, \dots, e_n)$ in the statement of the Theorem.

Another Remark

What average payoffs can be achieved by subgame-perfect Nash equilibria when the discount factor is not "sufficiently close to one"?

One approach

One way to approach this question is to consider a fixed value of δ and determine the average payoffs that can be achieved if the players use trigger strategies that switch forever to the stage-game Nash equilibrium after any deviation.

Smaller values of δ (less patience) make a punishment that will begin next period less effective in deterring a deviation this period.

Nonetheless, the players typically can do better than simply repeating a stage-game Nash equilibrium.

Another approach

A second approach (Abreu 1988) is that the most effective way to deter a player from deviating is to threaten to administer the strongest credible punishment

I.e., threaten to respond to a deviation by playing the subgame-perfect Nash equilibrium of the infinitely repeated game that yields the lowest payoff of all such equilibria for the player who deviated.

In most games, switching forever to the stage-game Nash equilibrium is not the strongest credible punishment

Hence some average payoffs can be achieved using Abreu's approach that cannot be achieved using the trigger-strategy approach.

In the Prisoners' Dilemma, however, the stage-game Nash equilibrium yields the reservation payoffs (that is, $e_i = r_i$), so the two approaches are equivalent.

Proof of Friedman's Theorem

Let $(a_{e1}, \dots, a_{en})$ be the Nash equilibrium of G that yields the equilibrium payoffs $(e_1, \dots, e_n)$.

Let $(a_{x1}, \dots, a_{xn})$ be the collection of actions that yields the feasible payoffs $(x_1, \dots, x_n)$.

(The latter notation is only suggestive because it ignores the public randomizing device typically necessary to achieve arbitrary feasible payoffs.)

Consider the following trigger strategy for player i :

Play a_{xi} in the first stage.

In the t^{th} stage, if the outcome of all $t - 1$ preceding stages has been $(a_{x1}, \dots, a_{xn})$ then play a_{xi}

Otherwise, play a_{ei} .

If both players adopt this trigger strategy then the outcome of every stage of the infinitely repeated game will be $(a_{x1}, \dots, a_{xn})$, with (expected) payoffs $(x_1, \dots, x_n)$.

We first argue that if δ is close enough to one, then it is a Nash equilibrium of the repeated game for the players to adopt this strategy.

We then argue that such a Nash equilibrium is subgame-perfect.

Suppose that all the players other than player i have adopted this trigger strategy.

Since the others will play $(a_{e1}, \dots, a_{e,i-1}, a_{e,i+1}, \dots, a_{en})$ forever once one stage's outcome differs from $(a_{x1}, \dots, a_{xn})$...

...player i 's best response is to play a_{ei} forever once one stage's outcome differs from $(a_{x1}, \dots, a_{xn})$.

It remains to determine player i 's best response in the first stage, and in any stage such that all the preceding outcomes have been $(a_{x1}, \dots, a_{xn})$.

Let a_{di} be player i 's best deviation from $(a_{x1}, \dots, a_{xn})$.

That is, a_{di} solves

$$\max_{a_i \in A_i} u_i(a_{x1}, \dots, a_{x,i-1}, a_i, a_{x,i+1}, \dots, a_{xn})$$

(Since any deviation triggers the same response by the other players, the only deviation we need to consider is the most profitable one.)

Let d_i be i 's payoff from this deviation:

$$d_i = u_i(a_{x1}, \dots, a_{x,i-1}, a_{di}, a_{x,i+1}, \dots, a_{xn})$$

(Again, we ignore the role of the randomizing device: the best deviation and its payoff may depend on which pure strategies the randomizing device has prescribed.)

We have:

$$d_i \geq x_i = u_i(a_{x1}, \dots, a_{x,i-1}, a_{xi}, a_{x,i+1}, \dots, a_{xn}) > e_i = u_i(a_{e1}, \dots, a_{en}).$$

Playing a_{di} will yield a payoff of d_i at this stage...

...but will trigger $(a_{e1}, \dots, a_{e,i-1}, a_{e,i+1}, \dots, a_{en})$ by the other players forever after...

...to which the best response is a_{ei} by player i , so the payoff in every future stage will be e_i .

The present value of this sequence of payoffs is

$$d_i + \delta \cdot e_i + \delta^2 \cdot e_i + \dots = d_i + \frac{\delta}{1 - \delta} e_i$$

Alternatively, playing a_{xi} will yield a payoff of x_i this stage and will lead to exactly the same choice between a_{di} and a_{xi} in the next stage.

Let V_i denote the present value of the stage-game payoffs player i receives from making this choice optimally (now and every time it arises subsequently).

If playing a_{xi} is optimal, then

$$V_i = x_i + \delta V_i$$

or $V_i = x_i / (1 - \delta)$

If playing a_{di} is optimal, then

$$V_i = d_i + \frac{\delta}{1 - \delta} e_i$$

as derived previously.

(Assume that the randomizing device is serially uncorrelated. It then suffices to let d_i be the highest of the payoffs to player i 's best deviations from the various pure-strategy combinations prescribed by the randomizing device.)

So playing a_{xi} is optimal if and only if

$$\frac{x_i}{1-\delta} \geq d_i + \frac{\delta}{1-\delta} e_i$$

or

$$\delta \geq \frac{d_i - x_i}{d_i - e_i}$$

Thus, in the first stage, and in any stage such that all the preceding outcomes have been $(a_{x1}, \dots, a_{xn})$, player i 's optimal action (given that the other players have adopted the trigger strategy) is a_{xi} if and only if:

$$\delta \geq (d_i - x_i)/(d_i - e_i).$$

Combining this observation with the fact that i 's best response is to play a_{ei} forever once one stage's outcome differs from $(a_{x1}, \dots, a_{xn})$, we have that it is a Nash equilibrium for all the players to play the trigger strategy if and only if

$$\delta \geq \max_i \frac{d_i - x_i}{d_i - e_i}$$

Since $d_i \geq x_i > e_i$, it must be that $(d_i - x_i)/(d_i - e_i) < 1$ for every i , so the maximum of this fraction across all the players is also strictly less than one.

It remains to show that this Nash equilibrium is subgame perfect.

That is, the trigger strategies must constitute a Nash equilibrium in every subgame of $G(\infty, \delta)$.

Recall that every subgame of $G(\infty, \delta)$ is identical to $G(\infty, \delta)$ itself.

In the trigger-strategy Nash equilibrium, these subgames can be grouped into two classes:

- (i) subgames in which all the outcomes of earlier stages have been $(a_{x1}, \dots, a_{xn})$
- (ii) subgames in which the outcome of at least one earlier stage differs from $(a_{x1}, \dots, a_{xn})$.

If the players adopt the trigger strategy for the game as a whole, then

- (i) the players' strategies in a subgame in the first class are again the trigger strategy, which we have just shown to be a Nash equilibrium of the game as a whole
- (ii) the players' strategies in a subgame in the second class are simply to repeat the stage-game equilibrium $(a_{e1}, \dots, a_{en})$ forever, which is also a Nash equilibrium of the game as a whole.

Thus, the trigger-strategy Nash equilibrium of the infinitely repeated game is subgame-perfect.

Collusion between Cournot Duopolists

Friedman (1971) was the first to show that cooperation could be achieved in an infinitely repeated game by using trigger strategies that switch forever to the stage-game Nash equilibrium following any deviation.

The original application was to collusion in a Cournot oligopoly, as follows.

Static Cournot game:

Aggregate quantity on the market: $Q = q_1 + q_2$

Market-clearing price is $P(Q) = a - Q$, assuming $Q < a$.

Each firm has a marginal cost of c and no fixed costs.

Firms choose quantities simultaneously.

In the unique Nash equilibrium, each firm produces:

$$q_C = (a - c)/3$$

Equilibrium aggregate quantity $2(a - c)/3$ exceeds the monopoly quantity:

$$q_m \equiv (a - c)/2$$

Hence both firms would be better off if each produced half the monopoly quantity, $q_i = q_m/2$.

Consider the infinitely repeated game based on this Cournot stage game when both firms have the discount factor δ .

Trigger strategy:

Produce half the monopoly quantity, $q_m/2$, in the first period.

In the t^{th} period, produce $q_m/2$ if both firms have produced $q_m/2$ in each of the $t - 1$ previous periods

Otherwise, produce the Cournot quantity, q_C .

The profit to one firm when both produce $q_m/2$ is:

$$\frac{\pi_m}{2} = (a - c)^2/8$$

which we will denote by $\pi_m/2$.

The profit to one firm when both produce q_c is $\pi_c = (a - c)^2/9$

If firm i is going to produce $q_m/2$ this period then the quantity that maximizes firm j 's profit **this period** solves

$$\max_{q_j} \left(a - q_j - \frac{1}{2}q_m - c \right) q_j$$

The solution is $q_j = 3(a - c)/8$, with associated profit of $\pi_d = 9(a - c)^2/64$ ("d" for deviation).

Thus, it is a Nash equilibrium for both firms to play the trigger strategy given earlier provided that

$$\frac{1}{1 - \delta} \cdot \frac{1}{2} \pi_m \geq \pi_d + \frac{\delta}{1 - \delta} \cdot \pi_c \quad (2.3.2)$$

Substituting the values of π_m , π_d , and π_c into (2.3.2) yields $\delta \geq 9/17$.

For the same reasons as in the previous section, this Nash equilibrium is subgame-perfect.

We can also ask what the firms can achieve if $\delta < 9/17$.

First approach:

We first determine, for a given value of δ , the most-profitable quantity the firms can produce if they both play trigger strategies that switch forever to the Cournot quantity after any deviation.

We know that such trigger strategies cannot support a quantity as low as half the monopoly quantity

But for any value of δ it is a subgame-perfect Nash equilibrium simply to repeat the Cournot quantity forever.

Therefore, the most-profitable quantity that trigger strategies can support is between $q_m/2$ and q_c .

To compute this quantity, consider the following trigger strategy:

Produce q^* in the first period.

In the t^{th} period, produce q^* if both firms have produced q^* in each of the $t - 1$ previous periods

Otherwise, produce the Cournot quantity, q_C .

The profit to one firm if both play q^* is

$$\pi^* = (a - 2q^* - c)q^*$$

If firm i is going to produce q^* this period, then the quantity that maximizes firm j 's profit this period solves

$$\max_{q_j} (a - q_j - q^* - c) \cdot q_j$$

Solution:

$$q_j = (a - q^* - c)/2$$

Associated profit:

$$\pi_d = (a - q^* - c)^2/4$$

It is a Nash equilibrium for both firms to play the trigger strategy given above if and only if:

$$\frac{1}{1 - \delta} \cdot \pi^* \geq \pi_d + \frac{\delta}{1 - \delta} \cdot \pi_C$$

Solving the resulting quadratic in q^* :

$$q^* = \frac{9 - 5\delta}{3(9 - \delta)}(a - c)$$

This is the lowest value of q^* for which the trigger strategies given above are a SPNE.

This expression for q^* is monotonically decreasing in δ

It goes to $q_m/2$ as δ approaches 9/17

It goes to q_C as δ approaches zero

Second approach (strongest credible punishment – Abreu 1986):

We show that Abreu's approach can achieve the monopoly outcome in our model when $\delta = 1/2$ (which is less than 9/17).

Abreu (1986) applies this idea to Cournot models more general than ours using an arbitrary discount factor

Consider the following "two phase" (or "carrot-and-stick") strategy:

Produce half the monopoly quantity, $q_m/2$, in the first period.

In the t^{th} period, produce $q_m/2$ if both firms produced $q_m/2$ in period $t - 1$,
 produce $q_m/2$ if both firms produced x in period $t - 1$

Otherwise produce x .

This strategy involves a (one-period) punishment phase in which the firm produces x and a (potentially infinite) collusive phase in which the firm produces $q_m/2$.

If either firm deviates from the collusive phase, then the punishment phase begins.

If either firm deviates from the punishment phase, then the punishment phase begins again.

If neither firm deviates from the punishment phase, then the collusive phase begins again.

The profit to one firm if both produce x is $(a - 2x - c) \cdot x$

Denote this by $\pi(x)$.

Let $V(x)$ denote the present value of receiving $\pi(x)$ this period and half the monopoly profit forever after:

$$V(x) = \pi(x) + \frac{\delta}{1 - \delta} \cdot \frac{1}{2} \pi_m$$

If firm i is going to produce x this period, then the quantity that maximizes firm j 's profit this period solves

$$\max_{q_j} (a - q_j - x - c)q_j$$

Solution:

$$q_j = (a - x - c)/2$$

Associated profit:

$$\pi_{dp}(x) = (a - x - c)^2/4$$

(dp stands for deviation from the punishment)

If both firms play the two-phase strategy above, then the subgames in the infinitely repeated game can be grouped into two classes:

- (i) collusive subgames, in which the outcome of the previous period was either $(q_m/2, q_m/2)$ or (x, x)
- (ii) punishment subgames, in which the outcome of the previous period was neither $(q_m/2, q_m/2)$, nor (x, x) .

For it to be a subgame-perfect Nash equilibrium for both firms to play the two-phase strategy, it must be a Nash equilibrium to obey the strategy in each class of subgames.

In the collusive subgames, each firm must prefer to receive half the monopoly profit forever than to receive π_d this period and the **punishment present value $V(x)$** next period:

$$\frac{1}{1-\delta} \cdot \frac{1}{2} \pi_m \geq \pi_d + \delta \cdot V(x) \quad (2.3.3)$$

In the punishment subgames, each firm must prefer to administer the punishment than to receive π_{dp} this period and begin the punishment again next period:

$$V(x) \geq \pi_{dp}(x) + \delta V(x) \quad (2.3.4)$$

Substituting for $V(x)$ in (2.3.3) yields

$$\delta \left(\frac{1}{2} \pi_m - \pi(x) \right) \geq \pi_d - \frac{1}{2} \pi_m$$

That is, the gain this period from deviating must not exceed the discounted value of the loss next period from the punishment.

(Provided neither firm deviates from the punishment phase, there is no loss after next period, since the punishment ends and the firms return to the monopoly outcome, as though there had been no deviation.)

Likewise, (2.3.4) can be rewritten as

$$\delta \left(\frac{1}{2} \pi_m - \pi(x) \right) \geq \pi_{dp} - \pi(x)$$

with an analogous interpretation.

For $\delta = 1/2$, (2.3.3) is satisfied provided $x/(a-c)$ is not between $1/8$ and $3/8$, and (2.3.4) is satisfied if $x/(a-c)$ is between $3/10$ and $1/2$.

Thus, for $\delta = 1/2$, the two-phase strategy achieves the monopoly outcome as a subgameperfect Nash equilibrium provided that $3/8 \leq x/(a-c) \leq 1/2$.

Other Dynamic Oligopoly Models

There are many other models of dynamic oligopoly that enrich the simple model developed here.

Two classes of such models:

State-variable models

Imperfect-monitoring models.

Rotemberg and Saloner (1986, and Problem 2.14) study collusion over the business cycle by allowing the intercept of the demand function to fluctuate randomly across periods.

In each period, all firms observe that period's demand intercept before taking their actions for that period;

In other applications, the players could observe the realization of another state variable at the beginning of each period.

The incentive to deviate from a given strategy thus depends both on the value of demand this period and on the likely realizations of demand in future periods.

(Rotemberg and Saloner assume that demand is independent across periods, so the latter consideration is independent of the current value of demand, but later authors have relaxed this assumption.)

Green and Porter (1984) study collusion when deviations cannot be detected perfectly:

Rather than observing the other firms' quantity choices, each firm observes only the market-clearing price, which is buffeted by an unobservable shock each period.

In this setting, firms cannot tell whether a low market-clearing price occurred because one or more firms deviated or because there was an adverse shock.

Green and Porter examine trigger-price equilibria, in which any price below a critical level triggers a punishment period during which all firms play their Cournot quantities.

In equilibrium, no firm ever deviates.

Nonetheless, an especially bad shock can cause the price to fall below the critical level, triggering a punishment period.

Since punishments happen by accident, **infinite punishments of the kind considered in the trigger-strategy analysis in this section are not optimal.**

Two-phase strategies of the kind analyzed by Abreu might seem promising

Abreu, Pearce, and Stacchetti (1986) show that they can be optimal.