1 Infinite horizon games

1.1 Bargaining

- Bargaining is the archetype of an economic problem. Applications: trading, politics, strikes, financial agreements, family,... .

- Bargaining situation is characterized by a mutual interest of maximizing the size of the cake, and a conflicting interest of maximizing the size of the share of the cake.

- There are two bargainers 1 and 2, sharing a pie of size 1. The domain of allocations is

$$X = \{(x_1, x_2) \in \mathbb{R}^2_+ : x_1 + x_2 = 1\}$$

- Players consume the pie. Player i's payoff from consuming $x_i \in [0, 1]$ is x_i.

- Recall the Ultimatum Game, where player 1 makes a take-it-or-leave-it offer to 2. In the unique SPNE, player 1 makes the offer $(1, 0)$ and 2 accepts the offer.

- I.e., as player 2 cannot *commit* to any punishment strategy, player 1 has all the *bargaining power* which gives him all the surplus

- Note: The game assumes that players only care about their own material payoff, and leave all kinds of psychological, e.g. fairness considerations aside.

- The prediction would not change if 1 makes a *finitely* many successive offers of which the first accepted is implemented.

- Situation changes if dramatically if 1 the game is of *infinitely* many successive offers.

EXAMPLE (Infinite offers game, no discounting): Player 1 makes an offer to 2 at each period $t = 0, 1, ...$ until an offer is accepted. An accepted offer is implemented. If no offer is ever accepted, both players get 0.

Any partition of X forms an SPNE of the game. Construct a strategy 1 always offers $(x_1, x_2) \in X$, and player 2 accepts any (y_1, y_2) such that $y_2 \geq x_2$, and rejects other offers.

- Thus in infinite games, the fact that only 1 makes offers does not give 1 any extra bargaining power.

- It is instructive to look at an infinite case with *discounting*.

- Player i has a discount factor $\delta_i \in (0, 1)$ that reflects the relative loss if a payoff is delayed one period. I.e. the present value of payoff x_i that received at period t is $x_i \delta^t$.

- Discount factor can be interpreted as a discounted value of a one unit of "utility". If r is the interest rate for one period, then $\delta = (1+r)^{-1}$. Alternatively, $1-\delta$ can be interpreted as the probability that the game ends at this period with 0 payoffs.

- Discounting reflects *time consistent* choice behavior.

- What is relevant is the assumption that players are *inpatient*. The closer δ is to 1, the more patient a player is.

EXAMPLE (Infinite offers game, discounting): Player 1 makes an offer to 2 at each period $t = 0, 1, ...$ until an offer is accepted. If an offer (x_1, x_2) is accepted at period t, players get payoffs $x_1\delta_1^t$ and $x_2\delta_2^t$. If no offer is ever accepted, both players get 0.

Any SPNE implements $(1, 0)$ at period 0. For if not, suppose $M > 0$ is the supremum SPNE of the payoffs for 2. Take any SPNE outcome $x = (x_1, x_2)$ such that $x_2 > \delta_2 M$, and suppose x is implemented at stage t. Then there is $\varepsilon > 0$ such that offer $(x_1 + \varepsilon, x_2 - \varepsilon)$ is accepted by 2 at stage t. Thus deviation by 1 from offer x to $(x_1 + \varepsilon, x_2 - \varepsilon)$ at t is profitable, a contradiction.

- Thus 1 has all the bargaining power in the game with discounting, even if he is very patient. This shows that there is an important discontinuity in the model when moving from the discounted version to the nondiscounted.

1.2 Rubinstein's alternating offers model

- Players 1 and 2 bargain according to an alternating offers procedure: If player i's offer in X is rejected at period t, then player $j \neq i$ makes an offer in period $t+1$. If an offer is accepted, then it is implemented. Player 1 starts.

- Players have discount factors δ_1 and δ_2. If an offer (x_1, x_2) is implemented at period $t+1$, players 1 and 2 receive payoffs $x_1\delta_1^t$ and $x_2\delta_2^t$, respectively.

THOEOREM (Rubinstein, 1982): The alternating offers model has a unique SPNE, where 1 always offers $x^1 = (x_1^1, x_2^1)$ such that

$$x_1^1 = \frac{1-\delta_2}{1-\delta_1\delta_2}, \; x_2^1 = \frac{\delta_2(1-\delta_1)}{1-\delta_1\delta_2},$$

and 2 always offers $x^2 = (x_1^2, x_2^2)$ such that

$$x_2^2 = \frac{1-\delta_1}{1-\delta_1\delta_2}, \; x_1^2 = \frac{\delta_1(1-\delta_2)}{1-\delta_1\delta_2},$$

and where 1 always accepts an offer that gives him at least x_1^2 and 2 always accepts an offer that gives him at least x_2^1.

- To prove the Theorem, note that there is no unilateral and profitable one-time nor infinite deviation. Thus the construction forms an SPNE.

- To see that this is the unique SPNE, identify the supremum M_i and the infimum m_i of the SPNE payoffs to i at any subgame where i proses. Since x^1 and x^2 are SPNE payoffs,

$$\begin{aligned} m_1 &\leq x_1^1 \leq M_1, & (1) \\ m_2 &\leq x_2^2 \leq M_2. & (2) \end{aligned}$$

First note that $m_2 \geq 1 - \delta_1 M_1$ since otherwise there is a 2's SPNE payoff u_2 such that $u_2 < 1 - \delta_1 M_1$, implemented at stage t. But now 2's offer $(\delta_1 M_1 + \varepsilon, 1 - \delta_1 M_1 - \varepsilon)$ at t generates him payoff $1 - \delta_1 M_1 - \varepsilon$, which is higher than u_2 for small enough ε, and it accepted by 1 since $\delta_1 M_1 + \varepsilon$ is more than 1 could ever get by rejecting the offer, and proposing in the next period. This contradicts the assumption that offering u_2 at stage t forms an SPNE.

Second, note that $1 - M_1 \geq \delta_2 m_2$ since otherwise $1 < M_1 + \delta_2 m_2$. Now there is an SPNE $u = (u_1, u_2)$ such that $u_1 > M_1 - \varepsilon$ and $u_2 > \delta_2 m_2$, the latter stemming from the fact that 2 can guarantee at least $\delta_2 m_2$ by rejecting u. This leads to a contradiction with $u_1 + u_2 = 1$, for small enough ε.

Thus

$$\frac{1 - M_1}{\delta_2} \geq m_2 \geq 1 - \delta_1 M_1,$$

i.e.

$$\begin{aligned} M_1 &\leq \frac{1-\delta_2}{1-\delta_1\delta_2} = x_1^1, \\ m_1 &\leq \frac{\delta_2(1-\delta_1)}{1-\delta_1\delta_2} = x_2^1. \end{aligned}$$

With (1), this implies

$$\begin{aligned} M_1 &= x_1^1, \\ m_2 &= x_2^1, \end{aligned}$$

and similarly

$$\begin{aligned} M_2 &= x_2^2, \\ m_1 &= x_1^2. \end{aligned}$$

Since $M_1 + m_2 = m_1 + M_2 = 1$, trade takes place immediately without delays.

- Note that

$$x_1^1 - x_1^2 = \frac{(1-\delta_1)(1-\delta_2)}{1-\delta_1\delta_2} > 0.$$

  Making an offer gives bargaining power. However, as δ_1 and/or δ_2 tend to one, the difference vanishes.

- Also, increase in δ_i increases i's payoff and decreases j's payoff. Thus becoming more patient increases players bargaining power!

A generalization:

- Let player i's payoff from share x_i be $u_i(x_i)$, where u_i is an increasing, concave function.

- Assume $\delta_1 = \delta_2 = \delta$.

- Now the unique SPNE is characterized by two allocations $x^1 = (x_1^1, x_2^1)$ and $x^2 = (x_1^2, x_2^2)$ such that

$$\begin{aligned} \delta u_1(x_1^1) &= u_1(x_1^2), \\ \delta u_2(x_2^2) &= u_1(x_2^1). \end{aligned}$$

  x^1 is now the allocation offered by 1 and x^2 the allocation offered by 2.

- Th existence is easy. Uniqueness rests on the same logic as the linear case.

- Note that

$$u_1(x_1^1)u_2(x_2^1) = \frac{u_1(x_1^2)u_2(x_2^1)}{\delta} = u_1(x_1^2)u_2(x_2^2).$$

Thus x^1 and x^2 induce utilities that lie on the same hyperbola $\{(v_1, v_2) \in \mathbb{R}^2_{++} : v_1 v_2 = u_1(x_1^2) u_2(x_2^1) \delta^{-1}\}$.

- As δ tends to 1, x^1 and x^2 become closer. In the limit, the hyperbola which runs through $u(x^1)$ and $u(x^2)$ parallels the Pareto frontier of the induced utility set.

1.3 Nash's axiomatic model

- Take the family $\mathcal{U}$ of all convex, compact, and comprehensive subsets U of $\mathbb{R}^2_{++}$.* Let $P(U)$ be the Pareto set associated to U. Call U symmetric

*Comprehensiveness: $u \in U$ implies $v \in U$ if $u_i \leq v_i \leq 0$ for $i = 1, 2$.

if $(u_1, u_2) \in U$ implies $(u_2, u_1) \in U$. Associate each U a *solution* $f(U)$ such that

A1 (Pareto optimality) $f(U) \in P(U)$.

A2 (Symmetry) If U is symmetric, then $f_1(U) = f_2(U)$.

A3 (Scale invariance) $f(\alpha U) = \alpha f(U)$ for all $\alpha \in \mathbb{R}^2_{++}$.

A4 (Independence of irrelevant alternatives) If $f(U) \in U'$ and $U \subset U'$, then $f(U') = f(U)$.

THEOREM (Nash, 1952): There is a unique solution $f^N : U \to \mathbb{R}^2_+$ that satisfies properties A1-A4. This solution is defined by the property that

$$f_1^N(U) f_2^N(U) = \max\{u_1 u_2 : u \in U\}.$$

- Sufficiency is easy. To see the necessity, let $f(U)$ be any solution meeting. Take any $U \in \mathcal{U}$ and identify $f^N(U)$. Rescale U by multiplying it with $\alpha = (f_1^N(U)^{-1}, f_2^N(U)^{-1})$. Then $f^N(\alpha U) = (1,1)$. Identify a symmetric triangle V such that $U \subset V$ and $(1,1) \in P(V)$. Since the slope of the tangent of αU at $(1,1)$ is -1, such triangle exists. Now A1 and A2 imply that $f(V) = (1,1)$. By A4, $(1,1) = f(\alpha U)$. Thus $f(\alpha U) = f^N(\alpha U)$. By A3, $f(U) = f^N(U)$.

- The fact that the axiomatic theory of Nash and the noncooperative model by Rubinstein predict the same outcome of the bargaining process is re-

markable. This is one prominent reason for why the Nash bargaining solution f^N has a prominent role in economic theory and applications.

2 Repeated games

- Many if not most social activities take place in long term, without clear understanding when the interaction will eventually stop.

- Firms compete, political groups negotiate, individuals interact and so on in frameworks where "trust" and "social pressure" can be important.

- Long term relationships are the theme of the vast literature on repeated games. They focus on the issues of cooperation and punishment which are difficult to capture in a static or even finite framework.

- Many social scientists and philosophers view a repeated the right framework to study the functioning of a social system and how it has evolved.

- Rich strategic structure of a repeated game creates new analytical and conceptual challenges.

2.1 Set up

- The building block of a repeated game is the *stage game*, a strategic game $\langle I, (A_i), (u_i) \rangle$ with finite action space A.

- Repeating the game means that it will be played many perhaps infinitely times in a sequential order. Payoffs of the stage game materialize at each period.

- Recall that cooperation in strategic games is difficult to achieve since there are no possibilities to punish or reward players after their actions.

- In repeated games, this problem can be overcome. The strategy space becomes large and, in general, actions can be made contingent on past occurrences.

- However, repeating a game has only limited value if the game (cf. the chain-store paradox).

EXAMPLE (Finitely repeated Prisoner's Dilemma): Let prisoner's dilemma be repeated for T times. Is cooperation feasible? At stage T, both players choose D since they know they are not punished afterwards. Since stage T

decisions are independent of the history, at stage $T-1$ the players choose D since they know they are not punished afterwards. Applying backwards induction, no cooperation can be entertained at any stage.

- An infinitely repeated game requires each player to choose a strategy for each period $t = 0, ...,$ conditional the observed history. Since only observed histories matter, strategies cannot be made contingent of mixed strategies, only pure ones.

- Let H be the set of possible histories of a game that repeats $\langle I, (A_i), (u_i) \rangle$ infinitely many times. That is

$$H = \{(a^0, ...a^t) : a^k \in \times_i A, \text{ for all } t = 0, ...\}.$$

- An infinite game strategy $a_i : H \to A_i$ specified an action for i for each possible history of past occurrences.

- As only realized actions can be observed, an action today cannot be made contingent on players' past mixed strategies. New subtle issues turn up in the connection of mixed strategies. We assume *pure* strategies.

- We assume discounting with equal discount factor $\delta \in (0, 1)$. Given the play path $(a^t)_{t=0}^{\infty} = (a^0, ...a^t, ...)$, the present value of player i's payoff is

$$(1 - \delta) \sum_{t=0}^{\infty} \delta^t u(a^t).$$

  Due to normalization, this number is between 0 and 1.

- Note: Any stage game Nash equilibrium can be supported as an SPNE equilibrium of a repeated game. For a proof, let a^* be a Nash equilibrium of the stage game $\langle I, (A_i), (u_i) \rangle$. Construct a strategy $a(\cdot)$ such that $a_i(h) = a_i^*$ for every $h \in H$. To see that this forms an SPNE, note that i's payoff is

$$(1-\delta) \sum_{t=0}^{\infty} \delta^t u(a_i^*, a_{-i}^*)$$

and a one time deviation at stage $t = 0$ generates

$$(1-\delta)(u(a_i, a_{-i}^*) + \sum_{t=1}^{\infty} \delta^t u(a_i^*, a_{-i}^*)).$$

The latter is bigger only if

$$u(a_i, a_{-i}^*) > u(a_i^*, a_{-i}^*),$$

which contradicts the assumption that a^* is a stage game Nash equilibrium.

- However, more SPNE equilibria may be supported.

EXAMPLE (Repeated Prisoner's Dilemma): Stage game payoffs are determined as follows:

		2: C	2: D
1	C	$3, 3$	$0, 4$
	D	$4, 0$	$1, 1$

Can the players support (C, C) in an SPNE? Consider the following strategies:

Nash reversion: Choose C until the other player chooses D. If the other player chooses D, then choose D forever.

Tit-for-tat: Choose C if the other player's most recent choice was C, and D if the other player's most recent choice was D.

Note that Grim and Tit-for-tat generate the same equilibrium path if both players adhere the strategy: Namely C is played infinitely since neither player is the first to deviate. Any *one-time* deviation leads, under Grim, to infinite play of D no matter what, but under Tit-for-tat to infinite play of D since the other player plays D and the deviation was one-time.

For Nash reversion to be an SPNE, note first that since (D, D) forms a Nash equilibrium of the stage game, it also forms a SPNE after a deviation from (C, C). A deviation is not profitable if

$$3 + 3\delta + 3\delta^2 + ... \geq 4 + \delta + \delta^2 + ... \ ,$$

or

$$\frac{3}{1-\delta} \geq 4 + \frac{\delta}{1-\delta}.$$

Hence if $\delta \geq 1/3$, the Grim is an SPNE.

Tit-for-tat forms an SPNE if deviating once from C and then reverting back to C is not profitable. A deviation gives 4 but reverting back 0, since then the other player plays D against which one has to play C. Without deviation, one obtains 3 in both periods. Thus a deviation is not profitable if

$$3 + 3\delta \geq 4,$$

or if $\delta \geq 1/3$, then Tit-for-tat is an SPNE.

- With patient enough players, cooperation can be sustained in infinite games. Rational players think of their long term prospects rather than short term gains.

- Grim strategy is very sensitive for the other player making mistakes. On the contrary, tit-for-tat allows mistakes and is simple. Moreover, it has a simple behavioral interpretation: *reciprocity.*

- Thus there are many equilibria in the repeated games. An easy general lesson is that *any* stage game payoffs above a Nash equilibrium payoffs can be induced in SPNE with high enough discount factor, in *any* game, by using the Nash reversion strategies (Tit-for-tat is not applicable in every game).

General result

- To express a powerful result on the size of the equilibrium outcomes denote by m_i player i's minmax payoff of the stage game $\langle I, (A_i), (u_i) \rangle$, i.e.

$$m_i = \min_{a_{-i}} \max_{a_i} u_i(a_i, a_{-i}).$$

FOLK THEOREM (Maskin-Fudenberg): A stage game strategy a^* can be supported in SPNE for high enough δ if $u_i(a^*) > m_i$ for all $i \in I$.

- To sketch the proof, identify player i's punishment strategy $b_{-i} \in \times_{j \neq i} A_j$ such that $u_i(a_i, b_{-i}) \leq m_i$ for all $a_i \in A_i$. Then construct the following strategy for the game where the play depends on the current "phase" of the game:

  - Cooperative phase: All players play a^*.

  - (i, T)-punishment phase for $T \geq 0$: All $j \neq i$ play b_{-i}.

  There are transitions from one phase to another according to following principles:

- From the cooperative phase to phase (i, T), if i deviates from a^* to a_i', where T satisfies the property that

$$u_i(a_i', a_{-i}^*) + \sum_{t=1}^{T} \delta^t m_i \leq \sum_{t=0}^{T} \delta^t u_i(a^*). \tag{3}$$

- From phase (i, T) to phase $(i, T-1)$ if all $j \neq i$ adhere b_{-i}. If $T = 0$, then move to the cooperative phase.

- From phase (i, T) to phase (j, T') if $j \neq i$ deviates from b_{-i} in phase (i, T), where T' satisfies the property that

$$\max_{a_i} u_j(a_j, b_{-i,j}, a_i) + \sum_{t=1}^{T'} \delta^t m_j \leq \max_{a_i} \sum_{t=0}^{T} \delta^t u_j(b_{-i}, a_i) + \sum_{t=T+1}^{T'} \delta^t u_i(a^*). \tag{4}$$

Note that since

$$u_i(a^*) > m_i, \text{ for all } i \in I,$$

there is high enough δ such that the desired (i, T)'s can be identified.

That is, a deviation by i from cooperation triggers a punishment of i, and a deviation by j from the phase that punishes i triggers a phase that punishes j. By (3) and (4), any deviation from the proposed strategy cannot be profitable, given that the other players adhere their strategy. Thus the strategy forms an SPNE. Since the play starts with a^*, it is played forever.

- The Folk Theorem reveals that infinite games have, in general great many equilibria. This raises an equilibrium selection problem.

- Pareto criterion is particularly problematic. Why should one appeal to Pareto optimality in the beginning if one cannot appeal to it during a punishment phase?

- Can one induce Pareto optimal equilibria with punishments in the Pareto frontier? $\Rightarrow$ Renegotiation proof SPNE.

- The problem: Minmax strategies need not be Pareto optimal.

- In general, yes. Since the equilibrium set is *convex.*

- Pure strategies *convexity* the stage game payoff space by altering pure actions with the appropriate frequencies: Let $A = \{0, 1\}$ and $u_1(0) = 0$,

$u_1(1) = 1$, and $u_2(a) = 1 - u_1(a)$, for all $a \in A$. Then, for any $x \in [0, 1]$ there is $(a^t)_{t=0}^{\infty} \in A^{\infty}$ such that if $\delta > 1/2$,

$$\begin{aligned} x &= (1-\delta) \sum_{t=0}^{\infty} \delta^t u_1(a^t), \\ 1 - x &= (1-\delta) \sum_{t=0}^{\infty} \delta^t u_2(a^t). \end{aligned}$$

To see this, construct $(a^t)_{t=0}^{\infty}$ by using the following rule: for any $t = 0, \ldots$

$$a^t = \begin{cases} 1, & \text{if} \quad x - (1-\delta) \sum_{s=0}^{t-1} \delta^s a^s \geq 1 - \delta, \\ 0, & \text{if} \quad x - (1-\delta) \sum_{s=0}^{t-1} \delta^s a^s < 1 - \delta. \end{cases} \tag{5}$$

For this procedure *not* to produce the desired result, it must be true that

$$x > (1-\delta) \sum_{t=0}^{\infty} \delta^t a^t.$$

By procedure (5), there is T such that $a^T = 0$ and $a^t = 1$ for all $t > T$. Since $\delta > 1/2$, we have

$$x - (1-\delta)\sum_{t=0}^{T-1} \delta^t a^t > (1-\delta)\sum_{t=T+1}^{\infty} \delta^t = \delta > 1-\delta.$$

But this violates the assumption that procedure (5) is followed. I.e. any payoff that is a convex combination of the maximal payoffs $(1,0)$ and $(0,1)$ can be generated.

- More generally, the set of payoffs from a repeated game is a convex set. This allows punishing a deviant player with his least profitable Pareto optimal payoff.

Imperfect monitoring, Trigger strategies

- Consider the following game: There is a worker hired by a firm, engaged into a repeated interaction.

- The worker may put either a high effort (H) or a low effort (L) on his job. If he chooses H, then the firm faces a demand (D) for its products with probability $1 - \varepsilon$, and the firm gets a payoff 1. With probability ε there is no demand (ND) for the firms' product, irrespective of the worker's effort, in which case the firm's payoff is 0.

- The per period salary of the worker is 2 independently of realization of the demand of the period. Each period's hiring decision is made independently.

- The per period payoffs are

		Worker	
		H	L
Firm		$1-\varepsilon, 1$	$0, 2$

- Can the firm hire the worker for each period and encourage the worker always to put a high effort?

- Suppose this is possible. Then any observed low demand is interpreted as a mistake without affecting the future hiring decisions.

- But then the worker's *choice* L does not affect the firm's hiring decision and hence it is profitable for the worker to continuously choose L. This breaks down the cooperative SPNE.

- Thus in any SPNE, at least sometimes a low demand *must* trigger a punishment phase.

- Infinite punishment would induce cooperative choices, but since a mistake takes place with probability one in the long run, no cooperation could be sustained in the long run. How should one design the punishment to induce at least some degree of cooperation?

- Let the punishment phase continue for $T-1$ periods after observing a low demand.

- The cooperative payoff V_T to the worker satisfies the recursion formula

$$V_T = 1 + (1-\varepsilon)\delta V_T + \varepsilon\delta^T V_T.$$

Thus

$$V_T = \frac{1}{1-(1-\varepsilon)\delta - \varepsilon\delta^T}.$$

To induce high effort in all periods, it suffices to have

$$V_T \geq 2 + \delta^T V_T.$$

I.e.

$$V_T \geq \frac{2}{1-\delta^T}.$$

Solving for δ^T,

$$\delta^T \leq \frac{2(1-\varepsilon)\delta - 1}{1-2\varepsilon}.$$

Thus if δ is low enough ($1 > 2(1-\varepsilon)\delta$), then there is no way to induce the worker to put a high effort. For high δ there is a minimal $T \geq 1$ such that the condition is met. Higher discount rate implies longer punishments

- Thus in the absence of perfect monitoring, punishments cannot be avoided even if both the firm and the worker know that worker always puts a high effort.

- But a simple trigger strategy may not always be optimal. For example, if the signal is very uninformative (ε close to $1/2$), it will in general be more profitable for the players to punish only after evidence has accumulated.