FDPE microeconomic theory, spring 2014
Lecture notes 3
Pauli Murto

- We consider first a well known application of perfect information games: alternating offers bargaining
- Then we consider repeated games (with perfect monitoring)
- Both models fall into the category of multi-stage games with observed actions, so that we can use sub-game perfect equilibrium as the solution concept

1 Alternating offers bargaining

1.1 One stage game

- Start with the simplest case: One period ultimatum game
- Two players share a pie of size 1.
- Player 1 suggests a division $x \in (0, 1)$.
- Player 2 accepts or rejects.
- In the former case, 1 gets x and 2 gets $1 - x$. In the latter case, both get 0.
- Given any $x \in (0, 1)$, the strategy profile $\{a_1 = x, a_2 = (\text{accept iff } a_1 \leq x)\}$ is a Nash equilibrium. So, there are infinitely many Nash equilibria.
- But once player 1 has made an offer, the optimal strategy for 2 is to accept *any* offer $a_1 < 1$ and he is indifferent with accepting offer $a_1 = 1$.
- What can you say about subgame perfect equilibria?

1.2 Two stages

- Suppose next that after 2 rejects an offer, the roles are changed
- 2 makes an offer x for player 1 and if accepted, she gets $1 - x$ for herself
- What is the SPE of this two-round bargaining game?
- What if players are impatient and payoff in stage 2 is only worth $\delta_i < 1$ times the payoff in stage 1 for player i.
- What is the SPE of this game?

1.3 Generalization to longer horizons: Alternating offers bargaining

- The player that rejects an offer makes a counteroffer
- Players discount every round of delay by factor δ_i, $i = 1, 2$
- If there are T periods, we can solve for a SPE by backward induction.
- Suppose we are at the last period, and player 1 makes the offer. Then he should demand the whole pie $x = 1$ and player 2 should accept.
- Suppose we are at period $T-1$, where player 2 makes the offer. He knows that if his offer is not accepted, in the next period player 1 will demand everything. So he should offer the least amount that player 1 would accept, that is $x = \delta_1$.
- Similarly, at period $T-2$ player 1 should offer division $x = 1 - \delta_2 (1 - \delta_1)$, and so on
- Can you show that as $T \to \infty$, then the player i who starts offers

  $$x = \frac{1 - \delta_j}{1 - \delta_i \delta_j}$$

  in the first period, and this offer is accepted?
- Note that the more patient player is stronger.
- What if there is infinite horizon? With no end point (i.e. all rejected offers are followed by a new proposal), backward induction is not possible
- Use the concept of SPE
- Notice that the subgame starting after two rejections looks exactly the same as the original game
- Therefore the set of SPE also is the same for the two games
- The famous result proved by Rubinstein (1982) shows that the infinite horizon game also has a unique equilibrium

Theorem 1 *A subgame perfect equilibrium in the infinite horizon alternating offer bargaining game results in immediate acceptance. The* unique *subgame perfect equilibrium payoff for player 1 is*

$$v = \frac{1 - \delta_2}{1 - \delta_1 \delta_2}.$$

1.4 Proof of Rubinstein's result:

- The part on immediate acceptance is easy and left as an exercise.
- Calculate first the largest subgame perfect equilibrium payoff $\overline{v}$ for player 1 in the game.
- Denote by $v_2(2)$ the continuation payoff to player 2 following a rejection in $T = 1$. The largest payoff for 1 consistent with this continuation payoff to 2 is:

$$1 - \delta_2 v_2(2)$$

- Hence the maximal payoff to 1 results from the minimal $v_2(2)$.
- We also know that

$$v_2(2) = 1 - \delta_1 v_1(3)$$

- Hence $v_2(2)$ is minimized when $v_1(3)$ is maximized.
- Notice next that the game starting after two rejections is the same game as the original one. Hence $\overline{v}$ is also the maximal value for $v_1(3)$.
- Hence combining the equations, we have

$$\overline{v} = 1 - \delta_2(1 - \delta_1 \overline{v})$$

  And hence

$$\overline{v} = \frac{1 - \delta_2}{1 - \delta_1 \delta_2}.$$

- Denote by $\underline{v}$ the smallest subgame perfect equilibrium payoff to 1. The same argument goes through exchanging everywhere words minimal and maximal. Hence we have:

$$\underline{v} = 1 - \delta_2(1 - \delta_1 \underline{v})$$

  and

$$\underline{v} = \frac{1 - \delta_2}{1 - \delta_1 \delta_2}.$$

  And thus the result is proved.

1.5 Comments:

- If $\delta_1 = \delta_2 = \delta \to 1$, then the SPE payoff converges to 50-50 split
- This is a theory that explains bargaining power by patience
- Cannot explain why there is often delays in bargaining
- Hard to generalize to more than two players
- Must have perfectly divisible offers
- Sensitive to bargaining protocol
- This model is based on Rubinstein (1982), "Perfect equilibrium in a bargaining model", Econometrica 50.

2 Repeated games

- An important class of dynamic games
- We only give some basic results and intuitions, and restrict here to the case of perfect monitoring (i.e. both players observe perfectly each others' previous actions)
- An extensive text book treatment: Mailath and Samuelson (2006), "Repeated games and reputations: long-run relationships", Oxford University Press
- In these games, the same "stage game" is repeated over and over again
- Player's payoff is most typically the discounted sum of the payoffs across stages
- The underlying idea: players may punish other players' deviations from nice behavior by their future play
- This may discipline behavior in the current period
- As a result, more cooperative behavior is possible

2.1 Stage game

- A stage game is a finite I-player simultaneous-move game
- Denote by A_i, $i = 1, ..., I$ the action spaces within a stage
- Stage-game payoff given by

$$g_i : A \to \mathbb{R}.$$

- In an infinite horizon repeated game, the same stage game is repeated forever

2.2 Strategies and payoffs

- Players observe each other's actions in previous periods
- Therefore, this is a multi-stage game with observed actions
- Denote by $a^t := (a_1^t, ..., a_I^t)$ the action profile in stage t
- As before, history at stage t, $h^t := \left(a^0, ..., a^{t-1}\right) \in H^t$, summarizes the actions taken in previous stages
- A pure strategy is a sequence of maps s_i^t from histories to actions
- A mixed (behavior) strategy σ_i is a sequence of maps from histories to probability distributions over actions:

$$\sigma_i^t : H^t \to \Delta(A_i).$$

- The payoffs are (normalized) discounted sum of stage payoffs:

$$u_i(\sigma) = \mathbb{E}_\sigma (1-\delta) \sum_{t=0}^{\infty} \delta^t g_i\left(\sigma^t\left(h^t\right)\right),$$

  where expectation is taken over possible infinite histories generated by σ

- The term $(1-\delta)$ just normalizes payoffs to "per-period" units
- Note that every period begins a proper subgame
- For any σ and h^t, we can compute the "continuation payoff" at the current stage:

$$\mathbb{E}_\sigma (1-\delta) \sum_{\tau=t}^{\infty} \delta^\tau g_i\left(\sigma^\tau\left(h^\tau\right)\right).$$

- A preliminary result:

Proposition 2 *If $\alpha^* = (\alpha_1^*, ..., \alpha_I^*) \in \Delta(S_1) \times ... \times \Delta(S_I)$ is a Nash equilibrium of the stage game, then the strategy profile*

$$\sigma_i^t\left(h^t\right) = \alpha_i^* \text{ for all } i \in I,\ h^t \in H^t,\ t = 0, 1, ...$$

is a sub-game perfect equilibrium of the repeated game. Moreover, if the stage game has m Nash equilibria $\left(\alpha^1, ..., \alpha^m\right)$, then for any map $j(t)$ from time periods to $\{1, ..., m\}$, there is a subgame perfect equilibrium

$$\sigma^t\left(h^t\right) = \alpha^{j(t)},$$

i.e. every player plays according to the stage-game equilibrium $\alpha^{j(t)}$ in stage t.

- Check that you understand why these strategies are sub-game perfect equilibria of the repeated game
- These equilibria are not very interesting. The point in analyzing repeated games is, of course, that more interesting equilibria exist too

2.3 Folk theorems

- What kind of payoffs can be supported in equilibrium?
- The main insight of the so-called folk theorems (various versions apply under different conditions) is that virtually any "feasible" and "individually rational" payoff profile can be enforced in an equilibrium, provided that discounting is sufficiently mild

Individually rational payoffs:

- What is the lowest payoff that player i's opponents can impose on i?
- Let
$$\underline{v}_i := \min_{\alpha_{-i}} \max_{\alpha_i} g_i \left(\alpha_i, \alpha_{-i}\right),$$
where $\alpha_i \in \Delta\left(S_i\right)$ and $\alpha_{-i} \in \times_{j \neq i} \Delta\left(S_j\right)$
- It is easy to prove the following:

Proposition 3 *Player i's payoff is at least $\underline{v}_i$ in any Nash equilibrium of the repeated game, regardless of the level of the discount factor.*

- Hence, we call $\{(v_1, ..., v_I) : v_i \geq \underline{v}_i \text{ for all } i\}$ the set of individually rational payoffs.

Feasible payoffs:

- We want to identify the set of all payoff vectors that result from some feasible strategy profile
- With independent strategies, feasible payoff set is not necessarily convex (e.g. in battle of sexes, payoff $\left(\frac{3}{2}, \frac{3}{2}\right)$ can only be attained by correlated strategies)
- Also, with standard mixed strategies, deviations are not perfectly detected (only actions observed, not the actual strategies)
- But in repeated games, convex combinations can be attained by time-varying strategies (if discount factor is large)
- To sidestep this technical issue, we assume here that players can condition their actions on the outcome of a public randomization device in each period
- This allows correlated strategies, where deviations are publicly detected
- Then, the set of feasible payoffs is given by
$$V = co\left\{v : \exists a \in A \text{ such that } g\left(a\right) = v\right\},$$
where co denotes convex hull operator

- Having defined individually rational and feasible payoffs, we may state the simplest Folk theorem:

Theorem 4 *For every feasible and strictly individually rational payoff vector v (i.e. an element of $\{v \in V : v_i > \underline{v}_i \text{ for all } i\}$), there exists a $\underline{\delta} < 1$ such that for all $\delta \in (\underline{\delta}, 1)$ there is a Nash equilibrium of the repeated game with payoffs v.*

- The proof idea is simple and important: construct strategies where all the players play the stage-game strategies that give payoffs v as long as no player has deviated from this strategy. As soon as one player deviates, other players turn to punishment strategies that "minmax" the deviating player forever after.

- If the players are sufficiently pationt, any finite one-period gain from deviating is outweighed by the loss caused by the punishment, therefore strategies are best-responses (check the details).

- The problem with this theorem is that the Nash equilibrium constructed here is not necessarily sub-game perfect

- The reason is that punishment can be very costly, so once a deviation has occurred, it may not be optimal to carry out the punishment

- However, if the minmax payoff profile itself is a stage-game Nash equilibrium, then the equilibrium is sub-game perfect

- This is the case in repeated Prisoner's dilemma

- The question arises: using less costly punishements, can we generalize the conclusion of the theorem to sub-game perfect equilibria?

- Naturally, we can use some low-payoff stage-game Nash equilibrium profile as a punishment:

Theorem 5 *Let α^* be a stage-game Nash equilibrium with payoff profile e. Then, for any feasible payoff vector with $v_i > e_i$ for every i, there is a $\underline{\delta} < 1$ such that for all $\delta \in (\underline{\delta}, 1)$ there is a sub-game perfect Nash equilibrium of the repeated game with payoffs v.*

- The proof is easy and uses the same idea as in above theorem, except here one uses Nash equilibrium strategy profile α^* as the punishment to a deviation

- Because the play continues according to a Nash equilibrium even after deviation, this is a sub-game perfect equilibrium

- Note that the conclusion of Theorem 5 is weaker than in Theorem 4 in the sense that it only covers payoff profiles where each player gets more than in some stage-game Nash equilibrium

- Is it possible to extend the result to cover all individually rational and feasible payoff profiles?
- Fudenberg and Maskin (1986) show that the answer is positive: in fact, for any $v \in V$ such that $v_i > \underline{v}_i$ for all i, there is a SPE with payoffs v given that δ is high enough (given an additional dimensionality condition on the payoff set: the dimension of the set V equals the number of players)
- See Fudenberg-Tirole book or the original article for the construction

2.4 Structure of equilibria

- The various folk theorems show that repeated interaction makes cooperation feasible as $\delta \to 1$
- At the same time, they show that the standard equilibrium concepts do little to predict actual play in repated games: the proofs use just one strategy profile that works if δ is large enough
- The set of possible equilibria is large. Is there a systematic way to characterize behavior in equilibrium for a given fixed δ?
- What is the most effective way to punish deviations?
- At the outset, the problem is complicated because the set of potential strategy profiles is very large (what to do after all possible deviations...)
- Abreu (1988) shows that all subgame perfect equilibrium paths can be generated by simple strategy profiles
- "Simple" means that these profiles consists of $I+1$ equilibrium *paths*: the actual play path and I punishment paths.
- A path is just a sequence of action profiles
- This is a relatively simple object - does not contain description of players' behavior after deviations
- The idea is that a deviation is punished by switching to the worst subgame perfect equilibrium path for the deviator:
  - Take a path as a candidate for a subgame perfect equilibrium path. We want to define a simple strategy profile that is a SPE and supports this path.
  - Find the worst sub-game perfect equilibrium path for each player. These are used as "punishment" paths.
  - Define players' behavior: follow the default path as long as no player deviates.
  - If one player deviates, switch to the punishment path of the deviator.

- If there is another deviation from the punishment srategy, again switch to the equilibrium that punishes deviator.
  - By one step deviation principle, this is a sub-game perfect equilibrium that replicates the original equilibrium path (recall that one-step deviation principle works for infinite horizon games with discounting)
  - Note that once all players follow these strategies, there is no deviation and hence punishment are not used along equilibrium path

- For a formalization of this, see Abreu (1988): "On the theory of infinitely repeated games with discounting", Econometrica 56 (2).

2.5 Example: oligopoly

- Finding the worst possible SPE for each player, as the construction above requires, may be difficult
- However, for symmetric games, finding the worst *strongly symmetric* pure-strategy equilibrium is much easier
- A strategy profile is strongly symmetric, if for all histories h^t and all players i and j, we have $s_i(h^t) = s_j(h^t)$
- Following the same idea as in Abreu (1988), we can construct the best strongly symmetric equilibrium by finding the worst punishment paths in the class of strongly symmetric equilibria
- This works nicely in games where arbitrarily low stage-game payoffs may be induced by symmetric strategies
- As an example, consider a quantity setting oligopoly model (with continuum action spaces)
- This originates from Abreu (1986), "Extremal equilibria of oligopolistic supergames", Journal of Economic Theory (here adapted from Mailath-Samuelson book)
- There are n firms, producing homogeneous output with marginal cost $c < 1$
- Firms maximize discounted sum of stage payoffs with discount factor δ
- Given outputs $q_1, ..., q_n$, stage payoff of firm i is

$$u_i(q_1, ..., q_n) = q_i \left(\max\left\{ 1 - \sum_{j=1}^{n} q_j, 0 \right\} - c \right).$$

- The stage game has a unique symmetric Nash equilibrium

$$q_i^N = \frac{1-c}{n+1} := q^N,\ i = 1, ..., n$$

with stage payoffs

$$u_i\left(q_1^N, ..., q_n^N\right) = \left(\frac{1-c}{n+1}\right)^2.$$

- The symmetric output that maximizes joint profits is

$$q_i^m = \frac{1-c}{2n} := q^m$$

giving payoffs

$$u_i\left(q_1^m, ..., q_n^m\right) = \frac{1}{n}\left(\frac{1-c}{2}\right)^2.$$

- Note that in this model, one sub-game perfect equilibrium is trivially $s_i\left(h^t\right) = q^N$ for all i and h^t

- Therefore, if δ is high enough, optimal outputs are achieved by Nash-reversion strategies: play q^m as long as all the players do so, otherwise revert to playing q^N forever

- However, cooperation at lower discount rates is possible with more effective punishments as follows

- Let $\mu(q)$ denote a payoff with symmetric output profile $q_i = q$ for $i = 1, ..., n$:

$$\mu(q) = q\left(\max\{1 - nq, 0\} - c\right)$$

- Let $\mu^d(q)$ denote maximal "deviation payoff" for i when others produce q:

$$\begin{aligned}
\mu^d(q) &= \max_{q_1} u_1\left(q_1, q, ..., q\right) \\
&= \begin{cases} \frac{1}{4}\left(1-(n-1)q-c\right)^2 \text{ if } 1-(n-1)q-c \geq 0 \\ 0 \text{ otherwise} \end{cases}
\end{aligned}$$

- Note that $\mu(q)$ can be made arbitrarily low with high enough q, allowing severe punishments

- Also, $\mu^d(q)$ is decreasing in q and $\mu^d(q) = 0$ for q high enough

- Let v^* denote the worst payoff achievable in strongly symmetric equilibrium (can be shown as part of the construction that a strategy profile achieving this minimum payoff exists)

- Given this, the best payoff that can be achieved in SPE is obtained by every player choosing q^* given by

$$q^* = \arg\max_q \mu(q) \tag{1}$$

subject to

$$\mu(q) \geq (1-\delta)\mu^d(q) + \delta v^*, \tag{2}$$

where the inequality constraint ensures that playinng q^* (now and forever) is better than choosing the best deviation and obaining the worst SPE payoff from that point on

- How do we find v^*?

- The basic insight is that we can obtain v^* by using a "carrot-and-stick" punishment strategy with some "stick" output q^s and "carrot" output q^c

- According to such strategy, choose output q^s in the first period and thereafter play q^c in every period, unless any player deviates from this plan, which causes this prescription to be repeated from the beginning

- Intuitively: q^s leads to painfully low profits (stick), but it has to be suffered once in order for the play to resume to q^c

- To be a SPE, such a strategy must satisfy:

  1. Players don't have an incentive to deviate from "carrot":

$$\begin{aligned} \mu(q^c) &\geq (1-\delta)\mu^d(q^c) + \delta\left[(1-\delta)\mu(q^s) + \delta\mu(q^c)\right] \text{ or} \\ \mu^d(q^c) - \mu(q^c) &\leq \delta\left(\mu(q^c) - \mu(q^s)\right) \end{aligned} \tag{3}$$

  2. Players dont' have an incentive to deviate from "stick":

$$\mu^d(q^s) \leq (1-\delta)\mu(q^s) + \delta\mu(q^c) \tag{4}$$

- To find the *optimal* "carrot-and-stick" punishment, we can proceed as follows:

- First, guess that joint optimum q^m can be supported in SPE. If that is the case, then let $q^c = q^m$, and let q^s be the worst "stick" that the players still want to carry out (knowning that this restores play to q^m), ie solve q^s from

$$\mu^d(q^s) = (1-\delta)\mu(q^s) + \delta\mu(q^m).$$

- If

$$\mu^d(q^m) - \mu(q^m) \leq \delta\left(\mu(q^m) - \mu(q^s)\right),$$

then no player indeed wants to deviate from q^m, and this carrot-and-stick strategy works giving:

$$v^* = (1-\delta)\mu(q^s) + \delta\mu(q^m)$$

- However, if
$$\mu^d(q^m) - \mu(q^m) > \delta(\mu(q^m) - \mu(q^s)),$$
then the worst possible punishment is not severe enough, and q^m cannot be implemented

- Then we want to find the lowest $q^c > q^m$ for which there is some q^s such that (3) and (4) hold

- This task is accomplished by finding q^c and q^s that solve those two inequalities as "=" (both "incentive constraints" bind)

- Note that this algorithm gives us the solution to (1) - (2): $q^* = q^c$ and $v^* = (1-\delta)\mu(q^s) + \delta\mu(q^*)$

- Is something lost by restricting to strongly symmetric punishment strategies? If $v^* = 0$, then clearly there cannot be any better asymmetric punishments (every player guarantees zero by producing zero in every period). Then restricting to strongly symmetric strategies is without loss

- However, if $v^* > 0$, then one could improve by adopting asymmetric punishment strategies

- It can be shown that q^* and v^* are decreasing in discount factor δ, and corresponding stick output q^s is increasing in δ

- That is, higher discount factor improves the achievable stage-payoff by making feasible punishments more severe

- For a high enough discount factor, we have $v^* = 0$ and $q^* = q^m$