Repeated Games

People often engage in ongoing relationships, and the same firms often compete against each other over time.

For example, in a dynamic game of Cournot competition, firms might be choosing output not just once, but once every month or week. In other words, the game of Cournot competition is repeated many times.

For example, over the summer, the same group of suppliers compete to provide milk to school districts in Texas for the coming school year. There is a sequence of milk auctions, one for each school district. A single auction can be modeled as a game of Bertrand (price) competition, but this game is repeated once for each school district in Texas.

In a repeated game, a player can make her action at a given stage depend on the entire history of play in previous stages. Thus, she may attempt to reward or punish others. She may be willing to sacrifice some of her payoff in one stage in order to beneficially affect other players' behavior in future stages.

A *repeated game* is a normal-form game (called the stage game) that is played repeatedly over discrete time periods (period 1, period 2, ... period t, ...) by the same players. The number of periods (or stages) the game is played, T, can be any finite positive integer. T can also be infinity, representing a situation in which players interact perpetually over time.

For the stage game, we have a set of action profiles

$$A = A_1 \times A_2 \times ... \times A_n$$

where A_i is the set of actions (strategies for the stage game) that player i can take at each stage.

The payoff function for player i in the stage game is denoted by $u_i(a) = u_i(a_1, ..., a_n)$.

A player's payoff in the repeated game is the sum of her payoffs across all of the stages. (For infinitely repeated games, we will take the payoff to be the average of the payoffs across all of the stages, since the sum can be infinite.) Thus, for finite T, if we let a^t denote the action profile at stage t, player i's payoff is

$$\sum_{t=1}^{T} u_i(a^t).$$

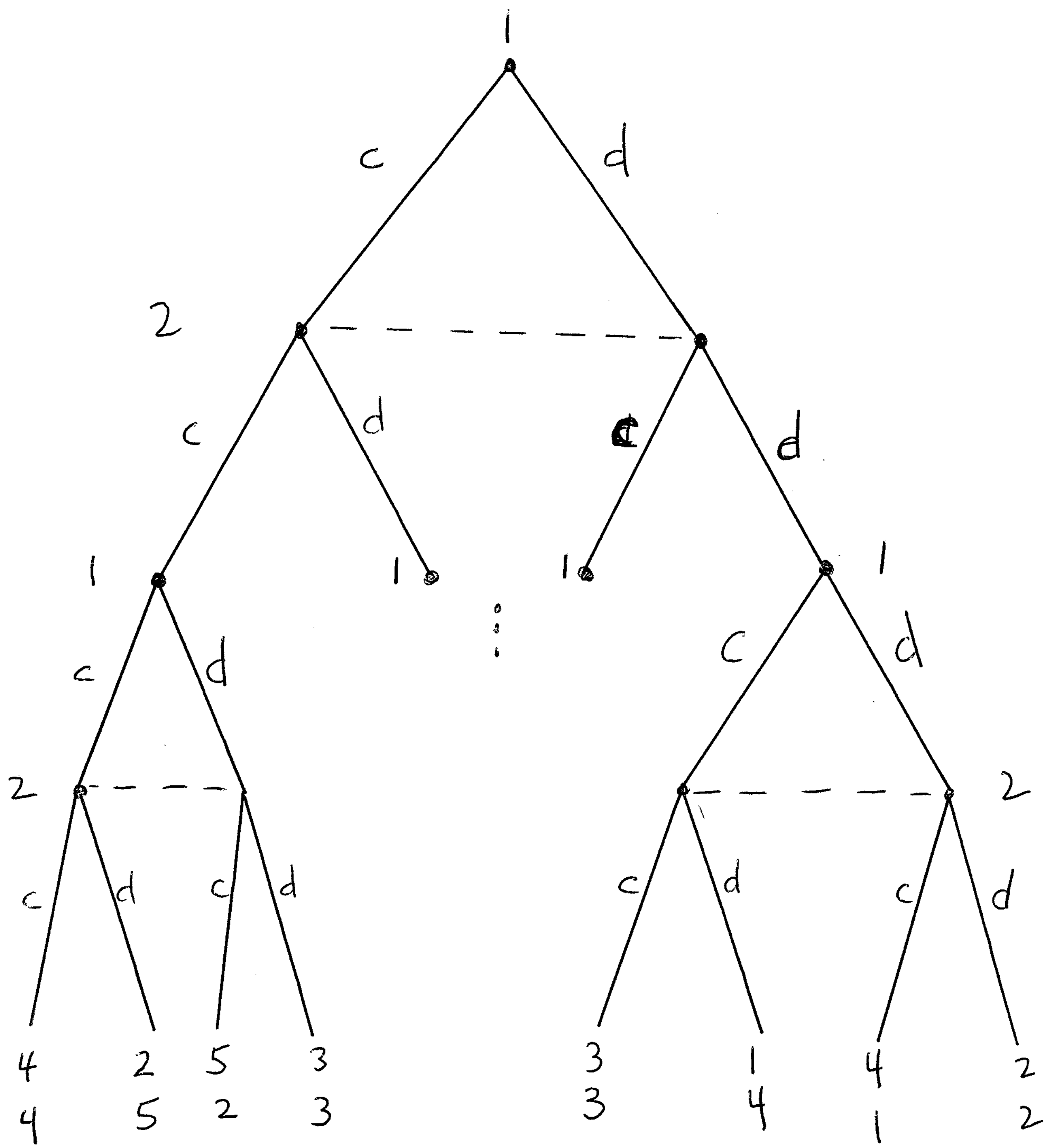
(Notice that I am using subscripts for players and superscripts for periods.)

Because players at stage t have observed the history of action profiles chosen in stages 1 to $t-1$, a *strategy* for player i in the repeated game specifies an action in A_i for each history.

In other words, a strategy specifies an action in period 1, an action in period 2 for each profile a^1, an action in period 3 for each history (a^1, a^2), an action in period t for each history $(a^1, a^2, ..., a^{t-1})$, and so on.

Notice that a repeated game is itself a game, satisfying all of the requirements in the definition. We can write the game in extensive form, although the game tree is usually too complicated to be useful. Each history of any given length defines a subgame of the repeated game.

[see the (unfinished) game tree for the twice repeated prisoner's dilemma game]



1
c
d
2
c
d
c
d
1
1
1
1
c
d
C
d
2
2
c
d
c
d
c
d
c
d
4 4
2 5
5 2
3 3
3 3
1 4
4 1
2 2

		player 2 cooperate	player 2 defect
player 1	cooperate	$2, 2$	$0, 3$
	defect	$3, 0$	$1, 1$

Notice that in the twice repeated prisoner's dilemma, the strategy of defecting in every period after every history is not a dominant strategy:

Suppose player 2's strategy is to cooperate in period 1, to cooperate in period 2 if player 1 cooperated in period 1, and to defect in period 2 if player 1 defected in period 1. (Choosing whatever action the other player chose in the previous period is called the "tit-for-tat" strategy.)

Then player 1's best response is to cooperate in period 1 and defect in period 2. According to player 2's strategy, she will then cooperate in period 2, yielding player 1 a payoff of 5. Always defecting only yields player 1 a payoff of 4.

Subgame Perfect Nash Equilibrium in Finitely Repeated Games

Let us find the SPNE of the T-period repeated prisoner's dilemma. After any history through the first $T-1$ periods, the profile chosen in period T must be a Nash equilibrium of the subgame corresponding to that history, so both players must be defecting in period T.

Working backwards, in period $T-1$, both players realize that their action cannot influence the profile chosen in the last period, so each player must be defecting in period $T-1$. (Cooperating leads to a lower payoff in period $T-1$, and does not change the fact that the only NE in the last period is (defect,defect)).

In period $T-2$, we have already argued that sequential rationality requires defection in all future periods, so the only sequentially rational choice is to defect in this period.

Thus, the game "unravels" and the only SPNE is for each player to defect after each history.

The only special feature of the prisoner's dilemma used to get this unraveling result is the fact that the stage game has a unique Nash equilibrium.

General Result: Consider any finitely repeated game where the stage game has a unique Nash equilibrium $(a_1^*, ..., a_n^*)$. Then the unique SPNE of the repeated game is for each player i to choose a_i^* after any history.

Note: It turns out that the finitely repeated prisoner's dilemma has many *Nash* equilibria, but all of them have an equilibrium *path* in which both players always defect. (If we think there can be a NE where someone cooperates on the equilibrium path, we see that the last person to do so is not best responding.)

Now suppose the following stage game with two pure strategy Nash equilibria is being repeated with $T = 2$.

		player 2		
		X	Y	Z
player 1	A	$4, 3$	$0, 0$	$1, 4$
	B	$0, 0$	$2, 1$	$0, 0$

The two NE of the stage game are (B, Y) and (A, Z). In the stage game, player 2's strategy X is never a best response.

Let us find the pure-strategy SPNE of the repeated game. In period 2, no matter what happened in period 1, the profile must be either (B, Y) or (A, Z).

		player 2		
		X	Y	Z
player 1	A	$4, 3$	$0, 0$	$1, 4$
	B	$0, 0$	$2, 1$	$0, 0$

Here is a SPNE (you can check):

player 1 chooses the strategy: A in period 1 and A in period 2 after every history.

player 2 chooses the strategy: Z in period 1 and Z in period 2 after every history.

Here is another SPNE:

player 1 chooses the strategy: A in period 1 and B in period 2 after every history.

player 2 chooses the strategy: Z in period 1 and Y in period 2 after every history.

If all players other than i are choosing a strategy that specifies an action in each period that does not depend on the history, then player i always has a best response that also does not depend on the history.

We can generalize the previous example, of playing one NE profile of the stage game in period 1 and another NE profile of the stage game in period 2.

General Result: In any repeated game, suppose we have a sequence of action profiles $(a^1, ..., a^T)$, where each a^t is a NE of the stage game. Then this sequence can be supported as the equilibrium path of a SPNE of the repeated game.

			player 2	
		X	Y	Z
player 1	A	$4, 3$	$0, 0$	$1, 4$
	B	$0, 0$	$2, 1$	$0, 0$

Back to our example, there are SPNE paths that are **not** just sequences of NE profiles of the stage game. Consider the following strategy profile:

player 1 chooses the strategy: A in period 1. In period 2, he chooses A if the period 1 profile was (A, X); otherwise, he chooses B.

player 2 chooses the strategy: X in period 1. In period 2, she chooses Z if the period 1 profile was (A, X); otherwise, she chooses Y.

This is a SPNE: No matter what happens in period 1, play in period 2 is a NE of the stage game. The profile for the full game is also a NE–Player 1 receives a payoff of 5, and by deviating to B, the payoff is at most 2. Player 2 receives a payoff of 7, and by deviating, the payoff is at most 5 (Z then Y).

		player 2		
		X	Y	Z
player 1	A	$4,3$	$0,0$	$1,4$
	B	$0,0$	$2,1$	$0,0$

Because the stage game has multiple NE, in the repeated game, a player can credibly commit to take an action that is not part of a NE of the stage game, based on the expectation that a (stage game) NE favorable to her will be played in period 2 if she takes the action, and a (stage game) NE unfavorable to her will be played in period 2 if she does not take the action.

In the T-period repeated game, there is a SPNE in which the path of the game is (A, X) for the first $T-1$ periods, then (A, Z) in period T. One interpretation is that there is a "social norm" that (A, X) is played until the last period, and (A, Z) is played in period T. If the social norm is violated, (B, Y) is played forever afterwards. This punishment threat incentivizes the social norm.

		player 2	
		cooperate	defect
player 1	cooperate	$2, 2$	$0, 3$
	defect	$3, 0$	$1, 1$

Recall that in the finitely repeated prisoner's dilemma, the players cannot maintain a reputation for cooperating in a SPNE, based on the threat of playing (defect,defect). Because the stage game has a unique NE, the game unravels.

Question–How reasonable is this unraveling argument when T is very large?

What about the **infinitely** repeated prisoner's dilemma? (Note: Watson assumes that payoffs are the discounted sum of stage game payoffs, while I look at the simpler and equally interesting case of average payoffs.)

Here is a SPNE:

player 1 and player 2 each choose the following strategy: Choose "cooperate" in period 1. After all histories in which both players have always cooperated, continue to cooperate. After all other histories, defect.

		player 2 cooperate	defect
player 1	cooperate	$2, 2$	$0, 3$
	defect	$3, 0$	$1, 1$

This strategy is called the "grim trigger" strategy, and when both players choose it, we have a SPNE. After any history of cooperation, continuing with the strategy yields an average payoff of 2, and defecting at any time yields an average payoff of 1. Thus, subgames following any history of cooperation are in a NE.

After any history with at least one "defect," continuing with the strategy (always defect) yields an average payoff of 1, and this is a best response to the other player's strategy (who will always defect). Thus, all subgames are in a NE.

The infinitely repeated prisoner's dilemma has many SPNE. For example,

1. Both players play a trigger strategy that is not "grim." Instead of lasting forever, a punishment stage lasts for K periods, and then the players choose "cooperate."

2. Both players choose "defect" after every history.

3. Both players choose the following strategy: Choose "cooperate" in period 1, and after all histories that alternate between (cooperate,cooperate) and (defect,defect), continue to alternate choices. After any departure from the alternating pattern, choose "defect."

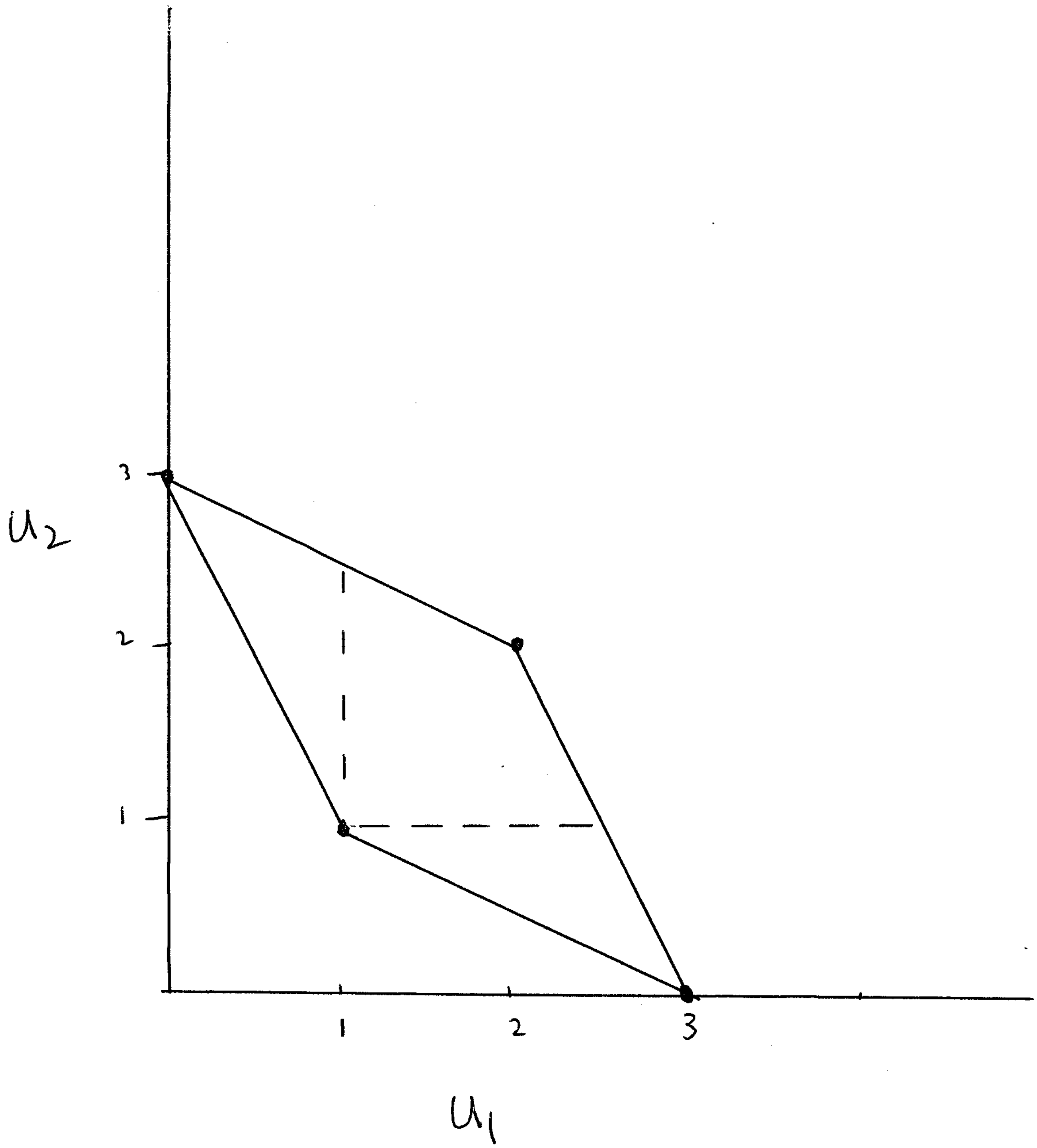
Notice that when we only observe the path of the game, say (cooperate,cooperate) in every period, we do not know whether the strategies form a NE or not.

Is the tit-for-tat strategy profile a SPNE? A NE?

		player 2	
		cooperate	defect
player 1	cooperate	$2, 2$	$0, 3$
	defect	$3, 0$	$1, 1$

The set of *feasible* average payoff profiles is the "convex hull" of the stage game payoffs. This refers to what is possible, without regard to whether the outcome is an equilibrium or not.

Of the feasible average payoff profiles, every profile that gives each player an average payoff of at least 1 is a SPNE payoff profile. Have the players alternate their actions in such a way as to give rise to the desired feasible payoff profile. If a player deviates from the pattern, both players forever choose "defect."



3
2
1
u_2
1
2
3
u_1

General Result (Folk Theorem): Consider any infinitely repeated game, and suppose that there is a Nash equilibrium of the stage game yielding a payoff profile $(w_1, w_2, ..., w_n)$. Then for any feasible average payoff profile $v = (v_1, v_2, ..., v_n)$ with $v_i \geq w_i$ for all $i = 1, ..., n$, there is a SPNE of the repeated game yielding the payoff profile v.

It is called a "Folk" theorem because many versions of this result were part of the conventional wisdom of economists and game theorists before it was formally proved.

Discussion of infinitely repeated Cournot and Bertrand competition: (i) cartel profits are available without explicit collusion, with implications for anti-trust, (ii) repeated interaction is predictably unpredictable, (iii) imperfect information and price wars.

Folk-Theorem type results are available in other dynamic situations that are not repeated games. For example, consider the infinite horizon Chain-Store game (infinite number of cities). This game has one long-lived player and an infinite number of players who only move once.

Here is a SPNE: The chain store fights whenever it is called upon to move after all histories in which it has always fought in the past; after all histories in which it has ever accommodated entry, the chain store accommodates.

Each potential entrant chooses "out" when the chain store has established a reputation for being tough by always fighting; if the chain store has ever accommodated entry in the past, a potential entrant will choose "in."

Why is this a SPNE?

To show that all subgames are in NE, consider a history in which it is a potential entrant's turn to move and the chain store has always fought when called upon to move. Then each potential entrant is receiving a payoff of 1, and no other strategy yields a higher payoff. The chain store is receiving an average payoff of 5, which is clearly a best response.

Consider a history in which it is the chain store's turn to move and the chain store has always fought when called upon to move. Then the chain store is best responding: if it fights it maintains its reputation and receives an average payoff of 5; if it accommodates it will face entry in all remaining cities and its average payoff is 2. The potential entrants are receiving a payoff of 1, and they cannot receive a higher payoff given the chain store's strategy.

Now consider a history in which the chain store has accommodated entry in the past. Then since the chain store will face entry in all remaining cities whatever it does, the chain store best responds by always accommodating entry. Since entry is being accommodated, the potential entrants are best responding by choosing "in."