Spontaneous Cooperation under Anarchy

1 International Cooperation

Cooperation, as you should recall, is part and parcel with our distributive games (the mixed motive games) between pure coordination and pure conflict. We have talked about bargaining in these situations and we have developed a large number of fairly useful concepts, such as commitments, credibility, private information, costly signaling, principal-agent relationships, preference aggregation, and various tactics to cope with problems related to all of these.

We now want to look a bit more closely at international cooperation when there is some degree of conflict. Unlike many of the situations that we examined, we shall assume that using force is not really a viable option despite anarchy. This is just common sense: would you threaten to invade with tanks if the other country levies a tariff on your bananas? (Well, maybe you would, but it is not common.)

We shall first examine the possibility of cooperation in distributive conflict situations arising "spontaneously" from self-interested behavior of the rational actors involved. We shall see that it is indeed possible to get order out of anarchy without any form of formal governance. However, we shall also find that there are significant costs to finding, creating, and maintaining this order. Given these costs, it may be impossible to overcome the problems of anarchy to ensure cooperation. We shall then see how institutions can arise to cope with these costs. In particular, we want to know when these rational actors would want to pay the costs of creating the institutions. Finally, we shall look at international organizations as the formal structures that support the institutions.

Let's begin with a very simple model of interaction between two actors. Some of you will recall the example of the Tariff Game from the first game theory lecture. This game, also commonly referred to as the Prisoner's Dilemma, models any situations where players would like to cooperate but have strong individual incentives to **free-ride** on the efforts of the other player. That is, they like the cooperative outcome but would rather avoid paying the costs of getting it if the other player is going to do it. Figure 1 gives a very simple formulation.

		Canada	
		C	D
U.S.	C	$1, 1$	$-1, 2$
	D	$2, -1$	$0, 0$

Figure 1: Provision of a Public Good: Pollution in the Great Lakes.

Suppose that the U.S. and Canada want to set environmental standards for limiting the pollution of the Great Lakes. The situation is such that even if one of them agrees to cease harmful emissions, the quality of the environment would improve sufficiently to sustain enough fish and recreation for both. However, limiting these emissions is costly, so each country would rather that have the other one do it. There is an additional problem that if a country implements the higher (costly) standards, its products manufactured in the plants where these standards are implemented become less competitive than products from polluting plants (that operate at lower costs).

Suppose that a sufficiently clean lake is worth \$2 billion to each country, and that limiting pollution costs \$1 billion. If both cooperate, the lakes are clean, and each gets the benefits net of cleanup costs, or \$1 billion. If both defect, neither pays the costs but neither enjoys the benefits, so payoffs of zero for both. If a player defects while the other cooperates, the lakes are sufficiently clean and the payoff is \$2 because the defecting player pays no costs. However, the "sucker" suffers both the costs and the additional revenue loss from not being competitive because of relatively expensive products (trade loss is \$2 billion). So its payoff is the benefit of clean lakes net of the cleanup and trade disadvantage costs, or $-\$1$.

We already know that the unique equilibrium of this game is (D, D), that is, both players defect. Given the incentive structure, defection is a strictly dominant strategy. The equilibrium outcome is dirty lakes, a bad result for both players. Under anarchy, players cannot credibly commit themselves to not defecting, and so they cannot enforce the cooperative outcome. Turning the game into one with sequential moves, and allowing either the U.S. or Canada to move first will not change the outcome. The unique subgame perfect equilibrium is still for both to defect.

This illustrates one of the most fundamental problems of international cooperation. In the absence of a way to enforce agreements, players would be stuck with suboptimal outcomes. Or would they? Is there a way for cooperation to emerge despite the commitment problem? That is, can players find ways of credibly committing themselves to cooperate even though there is no one to force them to live up to their promises?

2 Emergence of Cooperation with Repeated Interaction

We begin the answer to this question by noting that the U.S. and Canada are really not going anywhere in the foreseeable future. That is, they will have to play this game over and over again, for an indefinite period of time. Thus, the public good dilemma is an instance of **repeated interaction**, in which they have to engage for a very long time. In fact, nobody knows exactly how long.

We modify the model then in the following way. We divide time into discrete periods, each lasting some amount of time, for example, a year. We label all periods starting with the current one consecutively. Let t be the variable that denotes the period. So the first period is called $t = 0$, the next one $t = 1$, the next one $t = 2$, and so on.[1] In each period,

[1]There are technical reasons for labeling the first period with 0 instead of 1. The math is a lot easier to follow!

the two actors play the public good game and receive the payoffs from the outcome. So, if they both cooperate in $t = 0$, each gets 1, if Canada then defects while the U.S. cooperates, Canada gets 2 and the U.S. gets -1 in $t = 1$. Again, this continues indefinitely.

2.1 Time Preferences: Discounting

Now, common sense dictates that players value payoffs they receive today or in the near future more than payoffs they will receive in the distant future. You may care a lot about \$10 today but really that much about getting that \$10 ten years from now. The easiest way to represent this time-preference is with **discount factors**.

The discount factor, denoted by δ, is a number between 0 and 1, which measures the "patience" of the player. The higher the discount factor, the more patient the player is; that is, the more the player values the future. The lower the discount factor, the more impatient the players is; that is, the more the player privileges payoffs today to payoffs tomorrow.

Let me give you an example of how the discount factor works. Consider a player who is being offered \$10 today or a year from now. The **present value** from getting the money now is \$10, but the present value from getting it a year from now is $\delta \times 10$. That is, it is the value of the \$10 **discounted** by the discount factor δ. This measures how much getting the \$10 a year from now is worth to the player today. If $\delta = 0$, then the present value is \$0. The player really doesn't care at all about this \$10 a year from now; he is extremely impatient. If $\delta = 1$, then the present value is \$10. The player cares about this \$10 a year from now as much as about getting \$10 immediately; he is extremely patient. If $\delta = .8$, then the \$10 a year from now is worth only \$8 to him today. The player is somewhat patient, but would definitely prefer getting the money sooner than later. By varying δ between 0 and 1, we can represent different time-preferences for the same player or different players.

Suppose now that $\delta = .8$. Given the discount factor, we can evaluate the present value of receiving this \$10 at any time in the future. For example, if we want to know the present value of getting it two years from now, we only need to discount it twice, once for each year:

$$\delta \times \delta \times 10 = (0.8)^2 \times 10 = (0.64) \times 10 = 6.4$$

That is, it will be worth \$6.40 today (not surprisingly putting off the reward for another year diminishes its value, which was \$8 if put off for just one year). We follow the same procedure for three years:

$$\delta \times \delta \times \delta \times 10 = (0.8)^3 \times 10 = (0.512) \times 10 = 5.12$$

That is, it will be worth only \$5.12 today if received three years from now.

You can already see the general procedure here: to evaluate the present value of the payoff received some T periods in the future, you simply discount the value T times; that is, you multiply the payoff by δ^T. Nothing much to it. The one thing to remember is that for any $\delta < 1$, the present value gets smaller and smaller the further in the future you go. For example, for $T = 10$, the present value of \$10 will be $(0.8)^{10} \times 10 \approx 1.07$, or roughly \$1.07. For $T = 20$, it will be $(0.8)^{20} \times 10 \approx 0.12$, or 12 cents. Eventually it will be worth very close to zero.

Let's now put repeated interaction and discounting together. Suppose that the actors play the game for five years. This generates a sequence of outcomes, one for each year, depending on how they played in that year. One possible sequence is

$$(C,C),(D,C),(D,D),(D,D),(D,D)$$

That is, both began by cooperating in the first year. The U.S. then defected in the next year although Canada cooperated. Canada then retaliated for the defection by defecting as well, and so neither cooperated in the next three years. This sequence generates the following payoffs:

Period	1	2	3	4	5
t	0	1	2	3	4
U.S.	1	2	0	0	0
Canada	1	-1	0	0	0

Calculating the payoffs for both players is now simple. Assuming $\delta = 0.8$, they are

$$\text{U.S.}: \delta^0(1) + \delta^1(2) + \delta^2(0) + \delta^3(0) + \delta^4(0) = 1 + (0.8)(2) = 2.6$$
$$\text{Canada}: \delta^0(1) + \delta^1(-1) + \delta^2(0) + \delta^3(0) + \delta^4(0) = 1 + (0.8)(-1) = 0.2$$

That is, if the actors play the strategies that produce the five outcomes above, the present values of the stream of payoffs are \$2.6 billion for the U.S. and \$0.2 billion for Canada. Different strategies produce different outcomes, and so different streams of payoffs with different present values.

We are interested in analyzing situations that can last an indefinite number of periods. How do we do that? We can use the property of discounting which brings the present value of payoffs arbitrarily close to zero the further in the future you get them. We can analyze games that last an infinite number of periods. The nice thing about the discounting is that after a certain point, you "stop caring" about the payoffs, so assuming an infinite number of periods makes sense.

Now, if we assume that the game above lasts an infinite number of periods, we need to calculate infinite sums of discounted per-period payoffs. This seems pretty hard to do, but it turns out that it's not that bad. In particular, if a is some constant payoff that you get each period, the present value of getting this payoff an infinite number of periods when your discount factor is δ, is simply

$$\frac{a}{1-\delta}$$

That is, if you receive \$10 every year forever, the present value of this stream of payoffs for you, assuming $\delta = 0.8$, is $10/(1-0.8) = 10/0.2 = 50$. That is, this is worth \$50 to you. Of course, the formula for calculating the present value when you get different payoffs every period is messier, but we won't go into it. The simple version above will actually suffice to demonstrate very neat results of repeated interaction.

2.2 Playing Infinitely Repeated Games

If the game is infinitely repeated, you have an infinite number of strategies at your disposal even if the game itself has only two actions. For example, one strategy would be ALWAYS DEFECT, which prescribes that you defect in every period. Another simple strategy would be ALWAYS COOPERATE. Yet another would be "defect in odd periods, cooperate in even periods," or "cooperate always except every 10 periods," or whatever. You can clearly see that the possibilities are limitless.

This poses a problem for analysis: how do you find equilibria of games where players can play an infinite number of different strategies? Well, if you think about it, we're not really interested in finding all equilibria. We are interested in seeing whether particular equilibria exist. In our case, we want to know whether there are equilibria that can sustain cooperation indefinitely. If there are, what are the strategies that produce them?

So, instead of analyzing the repeated game for all possible equilibria (there's actually an infinite number of them), we look for "cooperative equilibria" and ask whether they exist and how to get them. We do this in the usual manner: we fix one player's strategy and look for best responses of the other player. If we find two strategies that are mutually best responses, we have an equilibrium.

Consider the strategy ALWAYS DEFECT (ALL-D). If U.S. plays ALL-D, what is Canada's best response? Since it cannot expect anything from the U.S., Canada's best response is ALL-D as well, to which the U.S. itself best-responds with ALL-D. That is, (ALL-D,ALL-D) is an equilibrium of the repeated game. In this equilibrium, both players receive zero. This is bad news. It appears that complicating the game did not help. Mutual defection is still an equilibrium!

Let's check another strategy: ALWAYS COOPERATE (ALL-C). Suppose U.S. plays ALL-C. Is ALL-C Canada's best response? Unfortunately not. If Canada plays ALL-C, it receives 1 in every period, and so its payoff is $1/(1-\delta)$. If it instead deviates to the strategy ALL-D, it will get 2 in every period, with a payoff $2/(1-\delta)$, which is always better regardless of the value of δ. So, deviation is profitable for Canada, so (ALL-C,ALL-C) is *not* an equilibrium. Things still look pretty bad.

One obvious problem with the above strategies that they are sort of stupid: they do not take into account past behavior. Essentially, they reduce the repeated game to its single play because they ignore the best part about repeated interaction: the ability to react to past actions of your opponent. A strategy that says "cooperate if your opponent has cooperated, defect otherwise" makes much sense. Generally then, we will be interested in **conditional strategies**, that is, strategies that condition current behavior on past actions.

Let's look at the strategy called GRIM TRIGGER (GRIM), which says "begin by cooperating and cooperate as long as both players have cooperated; in case of defection, switch to defecting and never cooperate again" (this is why the trigger—a defection— is called *grim*). Suppose both actors play this strategy. Is (GRIM,GRIM) an equilibrium? Let's check.

If U.S. plays GRIM, then following GRIM yields perpetual cooperation, and Canada gets a payoff of $1/(1-\delta)$. Suppose Canada defects, in which case U.S. switches to perpetual defection. Following its defection and given that U.S. now always defects, Canada will also choose to defect forever. So, the stream of payoffs for Canada will be 2 in the period it defects, and then 0 in all subsequent periods, which yields a total payoff of 2. Is deviation profitable? It will not be profitable if

$$
\begin{aligned}
2 &< \frac{1}{1-\delta} \\
2(1-\delta) &< 1 \\
\frac{1}{2} &< \delta
\end{aligned}
$$

In other words, if $\delta > 1/2$, sticking with GRIM will be more profitable than cheating given that U.S. plays GRIM. Suppose, however, that U.S. deviates and defects instead of cooperating. Would playing GRIM still be optimal?

If the U.S. defects, Canada would also defect, and since U.S. plays GRIM itself, the result will be perpetual defection. In this case, Canada gets the sucker payoff of -1 for one period, and then 0 forever after, giving it a total payoff of -1. But since the U.S. plays GRIM and always defects, Canada clearly cannot do better by cooperating. In other words, if the U.S. tries to cheat, Canada does best by punishing it.

This exhaust the possibilities to consider. Doing the equivalent calculations for U.S. yields the same condition for playing GRIM with Canada. We find a crucial result:

- If players are sufficiently patient, then (GRIM,GRIM) is a Nash equilibrium of the repeated game. The outcome is perpetual cooperation.[2]

We have reached an important conclusion: repeated interaction makes cooperation possible because it allows conditioning on past behavior. The reason GRIM can sustain cooperation is that it punishes defections by refusing to cooperate. When players care about the future, they care quite a bit about the losses they will incur if they defect, and so the short-term payoff from defection is outweighed by the long-term consequences of punishment.

Now the GRIM strategy is very fierce in its retaliation. Other conditional strategies have been proposed, the most famous of which is Tit-for-Tat (TFT), which begins by cooperating, and then does whatever the other player did in the previous period. That is, like GRIM, it rewards cooperation by cooperating, punishes defections by defecting, but unlike GRIM, it forgives as long as the other makes amends by cooperating. It is an intuitively-appealing strategy and it has actually been promoted by people as the best way to play the repeated game above because it supposedly can restore cooperation while preventing exploitation.

Unfortunately, some of the claims of TFT proponents are exaggerated. First of all, (TFT,TFT) is generally not a subgame perfect equilibrium. Intuitively, the problem is that it does not really restore cooperation after defection. To see this, suppose player 1 deviates from TFT for just one period and defects. Then player 2 responds in the next period by defecting. But since player 2 has cooperated in the period where player 1 defected, in the next period player 1 cooperates. So in that period player 1 cooperates (rewarding player 2's first period behavior), and player 2 defects (punishing player 1's defection in the first period). Obviously, players switch roles in the following period, where 2 cooperates but 1 defects. This actually generates a sequence of exploiter/sucker payoffs. But if player 2 did

[2]In fact, this equilibrium is also subgame perfect but it is trickier to show.

not retaliate in the first place, TFT would have restored cooperation immediately, and so not punishing is better, which implies TFT is not optimal. It is too-easily provoked and not forgiving enough.

However, it is possible to devise strategies that are *retaliatory* enough to deter defections by making them unprofitable, and yet *forgiving* enough to restore cooperation after the punishment is complete. Such strategies can sustain cooperation indefinitely for the same reason that GRIM can: they deter short-term exploitation by threatening longer-term punishment. In addition, they are more attractive than GRIM because they can restore cooperation eventually in case of defection.

This analysis is extremely useful because it makes it obvious that the emergence of cooperation critically depends on players being able to condition their behavior on past actions. This conditioning requires that players **monitor** the behavior of the other players. The digression about TFT also demonstrates that there exists a potential problem with enforcement. If the opponent plays TFT and defects, a player does better by actually forgiving that defection than retaliating. But if there is no retaliation, then the other player would simply always defect. We conclude that players must also find it in their interest to **enforce** punishments when necessary (which worked in GRIM).

- Cooperation possible under anarchy, can emerge in repeated interactions;
- Cooperation depends on monitoring and enforcement;
- There are still multiple equilibria, including non-cooperative ones.

As we saw earlier, (ALL-D,ALL-D) is an equilibrium of the repeated game (it is also subgame perfect). There are results in game theory, known as **Folk Theorems**, that demonstrate that *almost everything can happen in equilibrium in a repeated game*. That is, it is possible to sustain just about every level of cooperation, from perpetual defection, through intermittent cooperation, to full cooperation forever. In other words, the cooperative equilibria are only several among many, and there are equilibria with various degrees of cooperation. The point of the above exercise is to show that cooperation is possible under anarchy when interactions are repeated, not that it will, in fact, happen.

3 Summary

- Spontaneous cooperation under anarchy is possible with repeated interaction because repetition allows conditioning actions on past behavior. This makes punishment of defection possible, and can deter it when actors care about the future.
- Cooperation crucially depends on (a) valuation of future interaction, (b) monitoring of activity and verification of compliance, and (c) enforcement of punishment; that is, on information and credible commitments.
- Although there exist cooperative equilibria, they do so alongside non-cooperative ones, so there is no reason to expect that actors will coordinate expectations on the cooperative equilibria.