Institutions and International Organizations

We noted that many international interactions occur over long periods of time, and the single-shot models that we have examined so far do not really address this. To investigate how the shadow of the future affects behavior in the present, we extended our simple single-shot Prisoner's Dilemma (PD) game into a repeated PD game. In this version of the model, actors play the PD over time of indefinite duration.

To solve the iterated PD game, we introduced the concept of **time preferences**, which we modeled with **discount factors**. A discount factor is a measure of how highly the actor values the future relative to the present. We then extended the concept of strategy to include a complete description of how the entire iterated PD game should be played, not a single iteration of the per-period PD game. This resulted in an infinite number of possible strategies, but we focused on an important category of **conditional strategies**, where actors make their current and future actions contingent on what has transpired in the past.

Finally, we combined the ideas of repetition, discounting, and conditional strategies to find that that spontaneous cooperation under anarchy is possible but (a) it depends on players valuing the future sufficiently highly,[1] (b) it depends critically on players' abilities to monitor behavior to verify compliance, (c) it depends just as critically on their incentives to enforce the punishments, and (d) it is one outcome among many, and it is not clear how it is selected.

1 Transaction Costs

MONITORING. Ascertaining compliance can be very tricky. There is a lot of uncertainty, and so easy to disclaim responsibility for an outcome. For example, suppose U.S. observes more pollution than expected. It accuses Canada of having defected from the cooperative arrangement. Canada, however, claims full compliance and counters that the U.S. has failed to monitor its own producers properly, then produces new studies that show that pollution really is not up, and finally demonstrates that it is actually down along the Canadian shore, ending with an accusation that the U.S. has blatantly flaunted the agreement while Canada has been dutifully implementing all its provisions. Should the U.S. retaliate by defecting in the next period? Verifying compliance is a problem because it may involve substantial

[1] Note that in our exercise we simply found the lowest value of δ (the discount factor) such that the strategies (GRIM,GRIM) formed an equilibrium. We are not really interested in the precise value of the discount factor, just that it is possible to find at least one that works. In this case, any value of δ higher than $1/2$ would sustain cooperation in equilibrium. Generally, the discount factor can be approximated by various means empirically. For example, a government in a country with term limits will have a lower discount factor than a government that can potentially be re-elected many times.

costs (in this example, commissioning new studies, wasting time and effort to prove one's position right instead of doing something more productive, even monitoring own factories).

ENFORCEMENT is also costly, in this case because it involves suffering higher levels of pollution while punishing the other. Enforcement under anarchy usually takes the form of actions that are costly not only to the actor being punished but also to the one doing the punishment. Economic sanctions, for example, are not only detrimental to the target, but also to the sender, which is why very often they are not well implemented, which in turn shows why very often they fail to work at all. Although, as we have seen before the "hurt-more" criterion does not apply (that is, it does not matter who gets hurt more by the action, what's important is that the one being punished is hurt sufficiently relative to not being punished), when the action is costly, the credibility of the threat to carry it out is diminished.

Enforcement also has monitoring costs of its own: the punisher must verify compliance with its requirements so it knows whether to stop the punishment or continue it. This is also difficult: every target has incentives to claim it has complied. When the punishment involves more than one actor acting in coordination with the others (e.g. economic sanctions), there are also costs associated with making sure the others are cooperating in the punishment phase instead of secretly defecting, avoiding paying the costs, and free-riding on the punishment efforts of the others.

MULTIPLE EQUILIBRIA. The simple fact that there are more than one way to optimally play the game, and that each depends on players' expectations of each other's behavior poses at least a coordination problem in itself. We have already seen some of this in games of pure coordination where even minimal communication could, in principle, resolve the uncertainty. Of course, once we get into distributive conflict, this is no longer the case. Unless communication is costly, it will probably be ineffective in conveying information. So how do players coordinate their expectations?

We have seen that bargaining is a way to do this, and one can view coordination on some equilibrium as a result of a bargaining process between the players. Which particular equilibrium gets selected would then depend on the bargaining power of each player but in general unless the process yields something that is an equilibrium it would have been a waste of effort because at least one player would find it worthwhile to subvert the outcome by defecting from the arrangement. This is why sometimes an agreement may not reflect the interests of the conventionally stronger player: to ensure compliance in the future it may be necessary to settle for something that is not thoroughly disadvantageous to the other.

So bargaining can help coordinate expectations and select one of the many possible equilibria. However, bargaining is costly, sometimes very costly. It involves both time and effort that the players could productively spend elsewhere. Given the incentives to secure the best possible deal, the outcome may also be either no agreement or an agreement that is not an equilibrium, which usually means that players would have to renegotiate in the near future, spending all the time and effort again.

Generally, most international interactions involve a lot of effort that goes into organizing, negotiating, implementing, monitoring, and enforcing even the simplest agreements. The costs associated with these activities are called **transaction costs**. These costs provide

high barriers to cooperation. If it is difficult to monitor because it's too costly, conditional strategies may not work because it will not be clear whether the conditions for punishment have been met, and so whether retaliation must be triggered. Also, enforcement costs diminish the credibility of the threat to punish defections, which weakens deterrence, and so reduces the chances for cooperation. Finally, bargaining costs may prevent efficient coordination on a mutually beneficial equilibrium and players may end up selecting a suboptimal one.

Countries have interests in cooperating on the provision of a wide variety of public goods: financial stability, environmental crises, development assistance, arms control, and collective security. All these issues involve substantial transaction costs associated with locating a mutually acceptable bargain, implementing it, and then enforcing the agreement. In the anarchic international world then, players who are interested in cooperation will look for ways to facilitate its emergence by designing solutions that reduce the transaction costs involved. They do so by creating **organizations** whose rules and capabilities are such that they reduce transaction costs and implement the **institutions** that produce cooperation.

2 Institutions: Self-Enforcing Equilibria

Rational actors, who are aware of transaction costs and the fact that many of them are generated by the presence of asymmetric information, would naturally attempt to address them by creating institutions that mitigate uncertainty and otherwise reduce the transaction costs.

Institutions are enduring systems of social, humanly-designed, constraint on behavior. This constraint makes it possible for rational actors to credibly commit to complex actions, vastly expanding the scope of opportunities available to them. Institutions define the set of common expectations, through which actors achieve "order" in their interactions.

Usually, institutions are taken as granted, as part of the environment that determine what actions are physically available to the actors, or the information available to them. But while we do have to ask the important question of how institutions shape behavior, we need to ask the even more important ones of (a) how they are established in the first place, and (b) how are they supposed to function.

The central question is how institutional constraints can be maintained when actors are rational, and have at most only an instrumental interest in the institution for its own sake. That is, actors do not care about the institution itself but about the results it produces.

You should now be anticipating the answer from everything we have done thus far in the course: **Institutions themselves should be viewed as equilibria of rational behavior.** An institution is a type of equilibrium where an actor's behavior is conditioned on past behavior of others, and on its expectations about how other actors would react to its actions. That is, an institution is a long-lived equilibrium pattern of rational behavior in some underlying game that the actors are engaged in.

Institutions arise from the bargaining context between the participants, and its provisions must match the potential enforcement problems. **Enforcement of rules is endogenous**: that is, the rules are not taken for granted by arise out of the self-interested behavior

of the players. Notice how this requirement guarantees the credibility of the punishment threat, and therefore ensures that enforcement will work.

To understand an institution, we must examine how it establishes and maintains endogenous incentives for *cooperation* among rational selfish actors, which (from our foregoing discussion) means that we must examine how the manage to *coordinate* expectations in complex settings. As before, the role of *communication* is great, but that of *bargaining* is even more important.

Let me emphasize that the concept of institutions as equilibria is extremely far-reaching and has implications not only for international relations, but everyday social norms as well. If you get accustomed to thinking in terms of self-enforcing equilibria, you will find you understand a host of social phenomena far better than before.

We can distinguish between **informal** and **formal** institutions, depending on how centralized behavior is. Let's clarify the differences. Think about many players engaged in a repeated PD game. If each player simply plays GRIM with everyone else, then we do not have an institution in the usual sense of the term because there is no involvement of the group in whether individual actors cooperate or not. Each act of coordination benefits only the actor of the cooperating player, and each defection is punished only by the wronged player.

An informal norm of reciprocity would involve all players. For example, suppose that in each period, following the play of PD, each player sends a costly message to all other players except its current partner telling them whether the partner has cooperated or not. An actor reported to have defected inappropriately is then subject to subsequent punishment by all other players, who defect whenever they get to play with this actor until the actor makes amends by cooperating once unilaterally.

It is possible to show that strategies like these can support a cooperative equilibrium with many players, and that not only it is in the interest of each player not to defect, but, very importantly, it is also in their interest not to lie about what the partner has done.

As the number of players gets larger, there is an alternative institutional form that can generate even better payoffs for the actors. This requires centralized communication, and so it represents a formal institution, something that we shall call a *formal organization*. Suppose we designate a player to be the "director." Then, in each period each of the remaining players reports the identity of his partner to the director at some cost, and the director responds (for free) whether the partner is in good standing. Players then cooperate with partners reported to be in good standing and defect against partners in bad standing. Following the play, actors communicate (at cost) to the director whether their partner has defected inappropriately (that is, when not supposed to). Again, the only way to escape bad standing is to make a unilateral restitution.

It can be shown that these strategies also form an equilibrium in which players cooperate, do not lie about their partner's behavior, and where the director reports status honestly. Moreover, this equilibrium is easier to sustain when the number of players gets bigger, and the payoffs players get are higher than the informal one before (mostly because of the costliness associated with communication).

But we have just described a formal institution which resembles an organization we are likely to see. One player is designated to be the official who acts as the communication

hub for everyone else, but individual players are the ones who carry out the punishments. This finally leads us to international organizations, which from our perspective are simply formal institutions.

3 International Organizations: Implementing Institutions

Although creating an organization is costly itself, it may have lasting benefits that outweigh the costs of maintaining it. Organizations can reduce uncertainty by providing transparency (U.N. agencies, IMF, World Bank, OECD), defining standards (WTO rules, IMF conditionality, OPEC summits and quotas), reducing bargaining costs, helping establish credible commitments (cross-default clauses, Paris and London Clubs, NATO), and encouraging expectations of future interactions (U.N.).

Organizations are formal sets of rules and regulations that govern the interactions of their members. They help coordinate member expectations about each other's behavior. As we have seen many times already, behavior in strategic situations crucially depends on what one expects the other participants to do. These mutually consistent expectations may be difficult to form, especially when there are many potential ways of doing so, but they point to different answers. Essentially, one may think of organizations as providing focal points around which to coordinate expectations. Note how this helps with the third problem: equilibrium selection. Even if the cooperative equilibrium is one among many others, an organization can coordinate the expectations of its members such that everyone expects everyone else to expect the cooperative behavior, and so everyone cooperates.

One way to think about organization is as the formal physical entities that implement an institution. For example, the IMF implements the institution of international financial cooperation by (a) acting as a lender of last resort, (b) ensuring that member governments comply with its fiscal macroeconomic policies. How does enforcement work? First, the IMF distributes its aid in tranches, with subsequent disbursals conditional on government's performance. It is important to note that this conditionality work for small countries, like Poland and Bulgaria, but not for big ones, like Russia because (a) the U.S., who is the most important donor of the IMF has vested strategic interests in ensuring Russian stability, and (b) the Russian economy is too important for the region to allow it to go down the drain, both of which weaken the IMF's threat to punish noncompliance. As a result, the Russians have been getting money and mostly not implementing the reforms. The IMF has another lever, which is its "seal of approval" that it bestows on countries that meet their targets. This approval opens the doors for borrowing on the international markets. The private banks usually do not have enough resources to ascertain the likelihood of the debts being repaid (for which sound economic policies are necessary), so they use the IMF as an informational shortcut to determine how safe an investment would be. A country that is being blasted by the IMF is very unlikely to find alternate money to borrow on the private markets.

Generally, when you analyze an organization, you first should ask what institution it is supposed to implement. Then you ask how this institution (equilibrium) is supposed to work: What inducements are there to encourage cooperation, and what threats are there to

discourage defection? Most importantly, you should ask how these commitments are made credible; that is, how this institution is self-enforcing. If it is not, then the organization is likely to fail because it implements something that is not equilibrium behavior. In other words, it tries to force actors to behave against their interests without ensuring that the threat to punish them for misbehaving is credible.

One example of such failed organization is the predecessor of the United Nations, the League of Nations. It had several severe problems from the outset, the most important one being the fact that the U.S. was not a member. There were no adequate provisions for enforcement either, and countries relied mostly on economic sanctions, for which there were no adequate monitoring provisions to ensure that the countries supposed to punish were actually punishing (which is costly to them). In the end, the League failed to stop Italy from aggression in Africa against Abyssinia, then failed to stop Japan from aggression in Asia against China. It then fell apart.

The U.N.'s virtue in that respect is that it included the victors of the Second World War, and so the institution reflected the bargaining power of the most important states. One problem with the current setup is that this is no longer true. The Security Council does not take into account two extremely important and powerful states, Germany and Japan, or one may argue that it perhaps takes too many interests into account given that the U.S. is disproportionately stronger than everyone else. This brings me to the concluding remark for today: is the U.N. going to be irrelevant by rubber-stamping everything the U.S. demands, or is it going to be irrelevant by opposing it, and then getting ignored?

4 Summary

- **Transaction costs** are the effort, time, and resources spent for negotiating, implementing, monitoring, and enforcing agreements. They make cooperation more difficult, and even may preclude actors from coordinating on a beneficial equilibrium.
- Rational actors respond to the presence of these costs by creating **institutions** that mitigate uncertainty by coordinating expectations.
- Institutions must be viewed as **equilibria of rational behavior**, which means that all enforcement of their rules must be **endogenous**.
- Institutions are **informal**, where communication is decentralized, and **formal**, where communication is centralized. The more actors involved, the costlier communication becomes, and so the harder it may be to sustain informal institution.
- **International organizations** are formal institutions with relatively centralized communication, where the organization provides information about the members, but actors carry out the punishment themselves when one is called for.
- To understand any international organization, we must understand the bargaining context that produced the agreement, and how the rules it implements are endogenous. If the organization does not implement **self-enforcing rules**, or fails to reflect the actors' bargaining power, it will not work.