Coalition Formation in International Environmental Agreements and the Role of Institutions

Giulio Ecchia* and Marco Mariotti**

September, 22, 1997

Abstract: The aim of this paper is to discuss the role played by international institutions in achieving effective International Environmental Agreements. We emphasise the strategic nature of environmental negotiations and use a game theoretic model of coalitional bargaining to illustrate the main issues. We argue that international institutions can intervene in the framing of the strategic interactions between countries (i.e. setting the rules of the negotiation game) and can influence the actual agreement reached when different outcomes of the negotiation game can be equilibria.

J.E.L. Classification codes: Q3, C7

Keywords: coalition formation, environmental agreements, institutions, bargaining, cooperation

We would like to thank C.Carraro, H.Tulkens and F.Moriconi for comments on an earlier version. There have also been very helpful discussions on related topics with S.Brams. Financial support from Fondazione ENI Enrico Mattei is gratefully acknowledged.

*Department of Economics, University of Bologna, Bologna, Italy
**Department of Economics, Royal Holloway College, University of London, Egham, Surrey, U.K.

Corresponding author:

Giulio Ecchia
Department of Economics-University of Bologna
Piazza Scaravilli, 2 - 40126 Bologna, Italy
E-mail: ECCHIA@ECONOMIA.UNIBO.IT

Non technical summary

Many environmental problems - such as controlling climate change and emissions of greenhouse gases, the protection of the ozone layer and of biodiversity - are global, in the sense that they cannot be tackled effectively in isolation by any country and that they require international coordination of environmental policies. This has provided a strong impulse to organising international forums for negotiations between a large number of participants and to subscribing International Environmental Agreements (IEAs) that regulate the exploitation of the environment. But two main factors severely limit the effectiveness of such negotiations and agreements. The first factor is intrinsic to the strategic nature of the context. The gains from cooperation refer to the comparison between a situation of complete non-cooperation and a situation of complete cooperation: however, a country would in general prefer the situation in which it behaves non-cooperatively (thus avoiding whatever costs are involved in the implementation of 'green' policies) while the other countries cooperate amongst themselves (thus allowing the non-cooperating country to free-ride on the improved environment). The second obstacle to international cooperation is the lack of institutions with well-defined and effective enforcement powers. This often prevents IEAs from being binding, since the costs of non-compliance which can be imposed on the free-riders by the cooperating countries are too small, or too unlikely to occur, or too vaguely related to the breach (as may be the case for trade policy retaliations against environmental non-compliance). In such cases, only self-enforcing IEAs, which do not require any external authority, can be taken into consideration.

The aim of this paper is to discuss the role played by international *institutions* in achieving effective IEAs. We use the term 'institutions' broadly, to denote any international/supranational body which can be considered external to the specific environmental negotiation game being played by the countries. We focus on two issues for which institutions can be of importance: (1) *defining the rules* of the negotiation game. In fact, the self-enforcing nature, or lack thereof, of an IEA depends strongly on the exact features of the strategic situation at hand. Institutions can affect both the perceived incentives of participants (by directly modifying the game payoffs or by helping countries to *recognise* the nature of the game being played) and the structure of negotiations (for example, bilateral vs. multilateral talks); (2) *selecting* the most desirable agreement. Even when restricting attention to self-enforcing IEAs, there will typically be several of them. In other words, the negotiation game possesses multiple equilibria. The selection of one of them can only depend on elements which are not captured in the description of the game. Even if divested of enforcing power, an international institution can thus goad the countries towards 'focal' equilibria which are appealing.

Using three-country games in which each country can choose between a cooperative strategy (curbing emissions) or non-cooperative strategy (polluting), we explore the conditions which guarantee a stable coalition of countries which find it convenient to adhere to an IEA. By means of our analysis, we give some insights on how the intrinsically strategic nature of environmental negotiations calls for a relevant role of international institutions in effective IEAs. In particular, we argue that the nature of international agreements between non-myopic countries depends on the balance of credible threats that can support these agreements. For instance, we show that in situations that initially look unpromising for cooperation the very strength of the individual incentives to free-ride may indeed persuade countries to cooperate. In this respect, international institutions can intervene in the framing of the strategic interactions between countries (i.e. setting the rules of the negotiation game) and can influence the actual agreement reached when different outcomes of the negotiation game can be equilibria.

1. Introduction

Many environmental problems - such as controlling climate change and emissions of greenhouse gases, the protection of the ozone layer and of biodiversity - are global, in the sense that they cannot be tackled effectively in isolation by any country and that they require international coordination of environmental policies. This has provided a strong impulse to organising international forums for negotiations between a large number of participants and to subscribing International Environmental Agreements (IEAs) that regulate the exploitation of the environment. But two main factors severely limit the effectiveness of such negotiations and agreements. The first factor is intrinsic to the strategic nature of the context. The gains from cooperation refer to the comparison between a situation of complete non-cooperation and a situation of complete cooperation: however, a country would in general prefer the situation in which it behaves non-cooperatively (thus avoiding whatever costs are involved in the implementation of 'green' policies) while the other countries cooperate amongst themselves (thus allowing the non-cooperating country to free-ride on the improved environment). The second obstacle to international cooperation is the lack of institutions with well-defined and effective enforcement powers. This often prevents IEAs from being binding, since the costs of non-compliance which can be imposed on the free-riders by the cooperating countries are too small, or too unlikely to occur, or too vaguely related to the breach (as may be the case for trade policy retaliations against environmental non-compliance). In such cases, only self-enforcing IEAs, which do not require any external authority, can be taken into consideration.

The aim of this paper is to discuss the role played by international *institutions* in achieving effective IEAs. We use the term 'institutions' broadly, to denote any international/supranational body which can be considered external to the specific environmental negotiation game being played by the countries. We focus on two issues for which institutions can be of importance:

1. *Defining the rules* of the negotiation game. As we shall see, the self-enforcing nature, or lack thereof, of an IEA depends strongly on the exact features of the strategic situation at hand. Institutions can affect both the perceived incentives of participants (by directly modifying the game payoffs or by helping countries to *recognise* the nature of the game being played) and the structure of negotiations (for example, bilateral vs. multilateral talks).

2. *Selecting* the most desirable agreement. Even when restricting attention to self-enforcing IEAs, there will typically be several of them. In other words, the negotiation game possesses multiple equilibria. The selection of one of them can only depend on elements which are not captured in the description of the game. Even if divested of enforcing power, an international institution can thus goad the countries towards 'focal' equilibria which are appealing.

To analyse these issues, we employ a game-theoretic model of negotiations introduced in Mariotti (1997a) and used in Ecchia and Mariotti (1997) with specific reference to environmental negotiations. Here we will eschew all technicalities and refer to those papers for the details.

2. A Simple Model

The game can be described as follows. There is a set of countries or 'blocks' of countries, N. Each country may choose between a 'cooperative' strategy C (for 'Curbing' emissions) and a non-cooperative strategy P (for 'Pollute'). We assume that pollution on the part of other countries affects a country negatively irrespective of the behaviour of the country (polluting or cooperative) and of the identity of the polluters: so, the payoff for a country increases with the number of other countries who play cooperatively. A three-player game of this type is shown in Table 1.

	C	P	C	P
C	6,6,6	4,8,4	4,4,8	1,5,5
P	8,4,4	5,5,1	5,1,5	2,2,2

C P

Table 1: An environmental negotiation game

The negotiation procedure we imagine is the following[1]. At each stage, a strategy profile s {C,P} × {C,P} × {C,P} is the current status quo. Then, a *coalition* of countries S N may form and propose (or threaten) to deviate to a different set of strategies, that is, a different status quo. If somebody else, either a coalition of countries or an individual country, deviates, all member countries of S are free to propose to deviate further. The deviating coalition may be S itself: this means that there is no permanent commitment; agreements between member countries are no longer binding when the status quo changes. The countries are only interested in the payoff associated with a *permanent* status quo. So, the process continues in this way until there is a status quo from which nobody wishes to deviate. At this point the countries receive their payoffs[2]. For any coalition S, a *coalitional strategy* for S, specifies the proposal of coalition S at each possible status quo (thus we must distinguish carefully between a strategy (C or P) and a coalitional strategy). In order to define an appropriate notion of a self-enforcing agreement, we have to determine which coalitional strategies can be said to form an equilibrium, and which permanent status quo they will support. For the characterisation of a self-enforcing agreement we require the following properties. First, given a coalitional strategy profile, for a strategy profile to be an equilibrium (with respect to the given coalitional behaviour), no coalition should want to deviate from it once it is reached. Second, no coalition should be able to gain by switching to an alternative pattern of behaviour, that is, to an alternative coalitional strategy; and this must be true not only at the equilibrium status quo, but at

[1] Chandler and Tulkens (1997) provide an interesting but different formalisation of the coalition formation process with alternative hypotheses on the reaction of the non-members of a coalition when the coalition deviates.

[2] This procedure and the feature that only final payoffs matter is reminiscent of Brams' (1994) 'Theory of moves', in which a host of interesting problems of international politics (among many other topics) are analysed. One

any status quo. This imposes both a Nash and a Subgame Perfect equilibrium requirement, in that it rules out equilibria which are supported by countries making 'non-credible' threats or promises to pollute or to cooperate. The third requirement is that deviations from the current status quo should be 'motivated', in the sense that if a coalition departs from the current status quo, it must have some hope of ending up at a situation which is preferred to the status quo by all countries in the coalition. Finally, at a coalitional equilibrium point, no player can be forced below the payoff it can guarantee himself without the collaboration of any other player.

The previous discussion informally describes the notion of a *Coalitional Equilibrium* (Mariotti, 1997a). This concept aims at capturing some form of *farsightedness* on the part of the countries. When considering deviations from a current status-quo, any country or coalition of countries traces the indirect consequences of such a deviation, without naively considering the new status quo as the final one. This approach differs markedly from two other ways of modelling self-enforcing IEA's which have both proved popular in the analysis of environmental negotiations (see e.g. Hoel (1991), Carraro and Siniscalco (1992, 1993), Chandler and Tulkens (1993), Heal (1994), Barrett (1994a,b) and Carraro (1997)). The first relates to the adoption of some concept of *cartel stability* (D' Aspremont and Gabsewicz, 1986, Donsimoni *et al*., 1986), without however addressing the question of farsightedness. The second models environmental negotiations as *repeated games*. In that case, the attitudes of the players with respect to the passage of time play a major role, and, in addition, *actual* payoff-relevant deviations (and not only proposals) are crucial to support an equilibrium.

3. The Role of Institutions

---

distinguishing trait of the present model is its emphasis on *coalitions*, rather than on individual players as in Brams (1994).

In this section we use simple three-country examples to emphasise the role of institutions in international environmental negotiations[3]. The three-country assumption is not really restrictive for analysing situations with a *given* number of countries, as a wide variety of strategic situations can be generated even with this simplification. Here we limit ourselves to a four-way taxonomy. First, we distinguish between Prisoner's Dilemma (PD) and Chicken (CH) games. In a PD game, the non-cooperative strategy P is *dominant* for each country, while in a CH game the worst outcome for a country is the one in which everybody chooses P. The games in Table 1 and Table 2 are thus of the PD type. CH games are displayed in tables 3 and 4.

	C	P	C	P
C	3,3,3	1,5,1	1,1,5	0,4,4
P	5,1,1	4,4,0	4,0,4	2,2,2
	C		P	

Table 2: Strong Prisoner's Dilemma

	C	P	C	P
C	4,4,4	3,5,3	3,3,5	1,2,2
P	5,3,3	2,2,1	1,2,1	0,0,0
	C		P	

Table 3: Strong Chicken Game

	C	P	C	P
C	4,4,4	2,5,2	2,2,5	1,3,3
P	5,2,2	3,3,1	1,3,1	0,0,0
	C		P	

[3] Recall that for full analysis and proofs we refer to Ecchia and Mariotti (1997) and Mariotti (1997a).

Table 4: Weak Chicken Game

For each type, we further distinguish between a *strong* and a *weak* version. In the Weak PD, fixed the strategy played by one country, the game played by the remaining two countries is a two-player Prisoner's Dilemma (and hence the cooperative outcome Pareto-dominates the fully noncooperative one); while in the Strong PD, fixed the strategy of one country, the cooperative outcome is Pareto-dominated, in the two-player subgame played by the other two countries, by the fully noncooperative one. In the Strong CH game, each country only has an (individual) incentive to free-ride when everybody cooperates, but prefers to limit emissions when *at least* one other country is polluting. In the Weak CH, a country prefers the cooperative strategy only when *both* other countries are playing P, but if at least one other country behaves cooperatively, then it prefers to pollute. This terminology is chosen because the main feature of each type of game is emphasised in its strong version. So, for example, in the Weak PD there are at least limited individual incentives to cooperation, but not in the Strong PD. And while the 'disaster' character of a pollution situation occurs when even only one country pollutes in the Strong CH game, it takes two countries for this to occur in the Weak CH game.

When any of these games is played without communication among countries (in which case an appropriate solution concept could be the Nash Equilibrium), cooperation is either impossible (as in the PD games) or limited (as in the CH games). However, when the countries can communicate and thus the Coalitional Equilibrium concept is used, full cooperation (as well as partial cooperation between a subset of countries) becomes possible in equilibrium in all games, whereas complete noncooperation is not sustainable in equilibrium. In particular, in the strong PD game, only the fully cooperative outcome is possible in equilibrium, whereas in the three other games both full

cooperation and partial cooperation (between two countries) are possible in equilibrium. Not more than one country pollutes in equilibrium.

This contrast begins to provides broad support for the view that institutions and international forums for discussion can play a crucial role in achieving international cooperation in the use of environmental resources, even when these institutions are not endowed with enforcement powers. We need, however, to go deeper in analysis of the specific modes of institutional action.

We find that one of the most effective instruments in the hand of institutions with little or no enforcing power is the *framing* of a strategic situation in the most advantageous way to foster international cooperation. There are obvious limits for this type of action, but the point we are making is that, although in theoretical models payoffs must be exactly specified in advance, this is not the case in real negotiations, where the issues at stake are not always well defined and each country's perception of the payoff can be affected (see Young (1993) for a discussion of this point). For example, although a country cannot be persuaded that it would not benefit from free-riding on others (if they were not to react), the consequences of all-out non-cooperation may be less easy to compare with those of partial cooperation, in which case it may not be clear whether one is dealing with a CH or PD game[4]. Recall that experience shows that what determines payoffs is not only economic, readily quantifiable variables: disagreement on scientific facts and ideological barriers have proved to be major factors in real negotiations[5]. In addition, it is in the power of third parties to appropriately influence the set of issues which is the object of the negotiations. Indeed, the importance of issue-linking is one of the most emphasised in negotiation analysis (e.g. Sebenius, 1994).

---

[4] Greenberg (1995) provides an alternative framework, based on his Theory of Social Situations (Greenberg, 1990) in which cooperation is supported precisely *exploiting* the lack of a full specification of payoffs.

[5] This is very well documented for the case of the Rio Convention by the papers in Mintzer and Leonard (1994).

So, given this scope for action, how should a game be chosen? A surprising feature of our analysis is that the situation which at first sight appears to be the least promising, the Strong PD, is in fact the most conducive to cooperation. Indeed, the only Coalitional Equilibrium of this game is one where all countries cooperate. In the three other games, although full cooperation is possible in equilibrium, there are other less satisfactory equilibria. Beyond the limits of our particular formalisation, there seems to be a general lesson here, namely that it is crucial to distinguish between *myopic* incentives and *strategic* incentives for the countries when trying to achieve a self-enforcing IEA. The reason why full cooperation is the only possible equilibrium in the Strong PD is that the (myopic) incentives to pollute are so strong that the threat of retaliation by the other countries in case a country should attempt to free-ride is completely credible. Any temptation to free-ride will thus be chilled at the outset. In general, any country's strategic incentive to non-cooperation is limited by the incentives to non-cooperation of the other countries. A mediator aiming at fostering cooperation, therefore, should seek to frame the situation in such a way that it offers an appropriate *balance of threats*. A more subtle example of this feature occurs in the CH games. While in the PD games free-riding is immediately and in all cases vulnerable to a credible retaliation, it seems that in the CH games - where P is not always a dominant strategy - switching to P as a means of retaliation can be potentially harmful for the 'punishing' countries, because the most feared status quo is precisely one of full non-cooperation. But precisely because of the 'disaster' nature of this situation, it cannot be an equilibrium, thus making non-credible any counter-threat by a free-riding country not to depart from there[6].

[6] Note how in this reasoning we are making full use of the fact that every status quo is reversible. This is certainly not the case for some real situations of importance. However, the general point that a balance of threats has to sought retains its validity.

The scope for institutional action, however, is not limited to framing the payoffs in an appropriate way. The *structure* of negotiations will clearly play an important part. The specific equilibrium concept we use reflects one possible negotiation structure, in which all countries are free to negotiate with all others at any stage. In fact, especially with large number of countries, it may be practically difficult and costly to conduct such open negotiations, and limited talks between subsets of countries may be the only feasible way to proceed. In principle, this kind of limitation will affect negatively the possibility of cooperation, since analysis of our examples shows that global, Pareto-improving moves are used to support a cooperative equilibrium. The same examples, though, can be used to illustrate that in some cases a more practical institutional framework in which only coalitions of two countries are allowed to form can lead to satisfactory outcomes[7] (note carefully how this issue is distinct from the search for *limited agreements* between subsets of countries as opposed to complete cooperation, a point discussed by various authors. See e.g. the papers in Carraro, 1997). Consider the Strong PD. In that case, full cooperation can be obtained in equilibrium even when the grand coalition is prevented from moving. Then, the price for this limitation is that also a situation of complete non-cooperation becomes possible in equilibrium, which brings us to discuss the issue of *equilibrium selection*.

All of our example games, with the exception of the Strong PD, possess more than one (coalitional) equilibrium. As we have just mentioned, multiple equilibria emerge also in this case when different negotiation structures are considered. Which particular self-enforcing agreement countries will settle on is clearly a matter under the sphere of institutional control, since it remains unresolved on the basis of considerations of self-enforcement alone. A sophisticated institutional arrangement could thus incorporate a two-stage approach to moulding IEAs. In the first stage, the

[7] More precisely, one can mimic this framework by considering a modified Coalitional Equilibrium in which only

powers relating to the framing of payoffs and negotiation structure will be used to identify a set of possible equilibrium self-enforcing agreements, while in the second stage the question will be to focus the countries' attention on the most desirable ones. As the example of the Strong PD illustrates, this kind of two-stage approach may be useful for overcoming the practical impossibility of implementing an agreement which requires unwieldy 'grand' agreements between large number of countries.

4. Concluding Remarks

By means of our simple examples, we hope to have given some insights on how the intrinsically strategic nature of environmental negotiations calls for a relevant role of international institutions in effective IEAs. In particular, we have argued that the nature of international agreements between non-myopic countries depends on the balance of credible threats that can support these agreements. For instance, we have shown that in situations that initially look unpromising for cooperation the very strength of the individual incentives to free-ride may indeed persuade countries to cooperate. In this respect, international institutions can intervene in the framing of the strategic interactions between countries (i.e. setting the rules of the negotiation game) and can influence the actual agreement reached when different outcomes of the negotiation game can be equilibria.

We conclude with some brief observations regarding a line of research which could be interpreted as complementary to our analysis, and which relates to a positive analysis of the role of institutions in IEAs (see also Ulph (1997) and Carraro and Siniscalco (1997) on this point). First, as recently pointed out by Dixit (1996), the functioning of international institutions is often influenced by political and economic pressures by the member countries (loosely interpreted as transaction costs), which sometimes impair the effectiveness of institutional action. This poses a problem whenever either negotiation rules are defined (the issue is to find an agreement on rule setting) or

---

coalitions of a given maximum size can move. See Mariotti (1997b).

when more flexible courses of action are followed (the issue is then how much influence can negotiating countries exert to their individual or coalitional advantage). A second related point refers to the interplay between environmental and economic issues (such as trade issues) in international environmental negotiations. In this case, there is a potential overlapping between two or more different institutions, with an associated risk of coordination failure. Third, one should probably recognise that the term institution, as previously interpreted, is probably too generic, in the sense that it captures both supranational institutions and other organisations or groups, which can influence the outcome of negotiations (see e.g. Hurrel and Kinsbury, 1992). This recognition leads to a more explicit consideration of the role of interest groups in the negotiation process leading to an international agreement. In other words, there is a need for modelling more explicitly the political economy features of the negotiating process in order to evaluate its possible outcomes (with particular reference to the issue of equilibrium selection). Finally, even if we have considered self-enforcing agreements, it is clear that environmental treaties and protocols are revised over time in the light of new scientific evidence or of changed attitudes to environmental problems in the participant countries. In this respect, then, institutions can be particularly effective in the renegotiation of agreements, with particular reference to the balance of threats on which the new agreement can be reached.

References

Barrett,S., (1994a), "Self-Enforcing International Environmental Agreements", *Oxford Economic Papers,* **46**, 878-94.

Barrett,S. (1994b), "The Biodiversity Supergame", *Environmental and Resource Economics,* **4** , 878-94.

Brams,S., (1994), *Theory of Moves*, Cambridge, Cambridge University Press.

Carraro,C. (ed.), (1997), *International Environmental Negotiations: Strategic Policy Issues*, Cheltenham, U.K., Edward Elgar.

Carraro,C. and D. Siniscalco, (1992), "The International Dimensions of Environmental Policy", *European Economic Review,* **36**, 379-87.

Carraro,C. and D. Siniscalco, (1993), "Strategies for the International protection of the Environment", *Journal of Public Economics,* **52**, 309-28.

Carraro,C. and D. Siniscalco, (1997), "International Environmental Agreements: Political Economy Issues", *paper presented at the EEA 1997 meeting*, Toulouse.

Chandler,P. and H.Tulkens, (1993), "Strategically Stable Cost-Sharing in an Economic-Ecological Negotiations Process", in K.G. Maler, (ed.), *International Environmental Problems: Economic Perspective*, Dordrecht, Kluwer Academic Publishers.

Chandler,P. and H.Tulkens, (1997), "The Core of an Economy with Multilateral Environmental Externality", *International Journal of Game Theory*, **26**, 379-401.

D'Aspremont, C.A. and J.J. Gabsewicz, "On the Stability of Collusion", (1986) in G.F Matthewson and J.E. Stiglitz, (eds.), *New Developments in the Analysis of Market Structure*, New York, MacMillan Press.

Dixit,A.K., (1996), *The Making of Economic Policy: A Transaction-Cost Politics Perspective*, Cambridge, Mass., MIT Press.

Donsimoni, M.P., Economides,N.P. and H.M. Polemarchakis, (1986), "Stable Cartels", *International Economic Review,* **27**, 317-27.

Ecchia,G. and M.Mariotti, (1997), "The Stability of International Environmental Coalitions with Farsighted Countries: Some Theoretical Observations", in C. Carraro, (ed.), (1997), *International Environmental Negotiations: Strategic Policy Issues*, Cheltenham, U.K., Edward Elgar.

Greenberg, J., (1990), *The Theory of Social Situations: An Alternative Game-Theoretic Approach*, Cambridge, Cambridge University Press.

Greenberg, J., (1995), "Stable Incomplete Contracts in Dynamic Games", *mimeo*, McGill University.

Heal.G.,(1994), "The Formation of Environmental Coalitions" in C.Carraro, (ed.), *Trade, Innovation, Environment*, Dordrecht, Kluwer Academic Publishers.

Hoel, M., (1991), "Global Environmental Problems: The Effects of Unilateral Actions Taken by One Country", *Journal of Environmental Economics and Management*, **20**, 55-70.

Hurrel, A. and B. Kinsbury (eds.), (1992), *The International Politics of the Environment*, Oxford, Oxford University Press.

Mintzer, I.M. and J.A. Leonard, (eds.), (1994), *Negotiating Climate Change*, Cambridge, Cambridge University Press.

Mariotti, M., (1997a), "A Model of Agreements in Strategic Form Games", *Journal of Economic Theory*, **74**, 196-217.

Mariotti,M., (1997b), "Implementing Agreements in Restricted Talks: Some Examples", *mimeo*, Royal Holloway College, University of London.

Sebenius, J.K., (1994), "Towards a Winning Climate Coalition", in I.M. Mintzer and J.A. Leonard, (eds.), *Negotiating Climate Change*, Cambridge, Cambridge University Press.

Schelling, T.C., (1960), *The Strategy of Conflict*, Harvard, Harvard University Press.

Ulph, A., (1997), "Political Institutions and The Design of Environmental Policy in a Federal System with Asymmetric Information", *paper presented at the EEA 1997 meeting*, Toulouse.

Young, O.R., (1993), "Negotiating an International Climate Regime: The Institutional Bargaining for Environmental Governance" in N.Choucri, (ed.), *Global Accord: Environmental Challenges and International Responses*, Cambridge, Mass., Mit Press.