INTERNATIONAL ENVIRONMENTAL AGREEMENTS: A LITERATURE REVIEW

Alexis Ioannidis
University of Macedonia

Andreas Papandreou
University of Macedonia

Eftichios Sartzetakis
University College of the Cariboo

1.Introduction

Some of the most important environmental problems urgently calling for solution are problems related to transboundary pollution. Environmental problems such as ozone depletion, climate change and marine pollution have been the focus of intense negotiations at the international level over the past decade. International environmental agreements have been reached in some areas, such as the Montreal Protocol on substances that deplete the ozone layer and a number of agreements on marine pollution (e.g. the International Conference on Protection of the Marine Environment from Land-Based Sources, the International Convention for the Prevention of Pollution from Ships and the International Convention for the Prevention of Marine Pollution by Dumping of Wastes and Other Matter). In some other areas there is still an ongoing process of negotiations, such as the current negotiations following the UN Conferences on Environment and Development in Rio de Janeiro and Kyoto. Although there are a great number of international agreements, these agreements are characterized either by generality and vagueness in defining targets or by small number of signatories when precise targets are set. It is easy for countries to sign general statements expressing interest and willingness to improve their environmental behavior but it is extremely difficult to commit at achieving particular targets. The long and painful negotiations on the climate change problem also provide evidence of the difficulties associated with achieving large, stable and substantial international environmental agreements.

The formation and development of international environmental agreements has been the subject of a fast growing branch of the economic literature over the past decade, which though originates in the 1960s and 1970s. Recognizing the existing interdependencies between countries' choices, which lead to the observed strategic positioning of countries involved in negotiations of environmental agreements, a great part of this literature uses game theory as the tool of analysis. The purpose of the present work is to provide a critical review of this part of the literature on international environmental agreements. Our goal is to provide a systematic account of the issues examined, rather than an exhaustive presentation of the literature.

From the point of view of economic theory, environmental problems constitute one form of market failure –due to the absence of well defined property rights and markets-- and for that, some type of government intervention is needed for their resolution. Contrary to local environmental problems which can be addressed by existing governmental bodies, at the international level there is luck of a supra-national authority that could implement and enforce environmental policies on sovereign states. Thus, international agreements have to be designed so that they are self-enforcing.

There are two important design issues that self-enforcing international environmental agreements have to address. The first is that in order to be acceptable by a large number of countries the agreements have to be profitable to all countries. The profitability to all members of a coalition cannot be taken for granted especially when high asymmetry in both costs and benefits from reducing pollution exists among countries. Thus, in designing the agreement some type of profitability condition for all members has to be imposed.

Assuming that the agreement is designed to be profitable and countries have an incentive to join in, they also have incentives to cheat on the agreement ex-post. That is, countries have an incentive to forfeit their obligations of reducing their emissions so that they minimize their costs at the same time that they enjoy the benefits of lower emissions achieved by the rest of the countries. For some countries, the free-riding incentives may be large, undermining the stability of the coalition. Therefore, some type of stability constraint has to be incorporated into the design of the agreements so that they become stable.

The main body of the literature agrees that the number of signatories of self-enforcing, profitable and stable coalitions is very limited. Regardless of the degree of asymmetry among countries, there exist strong free-riding incentives that restrict the number of participants. This result seems to be robust both within a static and, what is more surprising, within a dynamic framework. However, it has been shown that the size of a stable coalition could increase substantially when the model is expanded in the following directions: (i) to account for non-material aspects in countries' benefits from joining the coalition; (ii) to account for interdependencies in countries' cost and benefit functions; and (iii) to allow countries to anticipate other countries' reactions to their choices instead of acting myopically.

Given that a stable, self-enforcing agreement will be signed by only few countries, which have an incentive to increase the number of participants, the natural next step is to consider ways in which the initial coalition can be expanded. The literature has identified two different types of instruments that can be used in expanding a coalition, namely direct transfers and issue linkages. A set of transfers can be derived to induce cooperation by all countries, that is, to make the agreement profitable for all countries. However, if transfers are not supported by some sort of commitment by at least the core countries, the agreement is subject to free riding and thus, not stable. Therefore, some sort of credible commitment is necessary to support side payments in expanding a stable coalition. The necessary degree of commitment varies depending on the assumptions of the model that is used.

Alternatively, cooperation in environmental issues can be linked to other issues affecting the negotiating countries. Two main issue linkages are examined in the literature, trade sanctions and R&D cooperation. It is shown that issue linkage substantially increases the size of an environmental coalition.

The rest of the paper is organised as follows. The basic framework for examining international environmental agreements is laid out in section 2. Section 3 presents the benchmark cases used in the literature as well as a general description of the self-enforcing international environmental agreements. Section 4 presents the main developments in the literature organised along three issues: (i) determining the size of the coalition, (ii) expanding the coalition and (iii) the choice of policy instruments. Conclusions are given in the final section.

2. The basic framework

This section presents the basic elements of the analysis of international environmental agreements. Because of the resulting complexity of the analysis, specification of production and utility function is avoided except when absolutely necessary in addressing certain questions. The majority of works reviewed here use a social welfare defined in very general terms. A number of assumptions are employed regarding the welfare function, which we denote as model assumptions. In defining the form of the game, both the cooperative and non-cooperative approaches are used in the literature within static, repeated and dynamic frameworks.

2.1 The objective function

At the first level, the link between economic activity and the physical environment is established in order to generate an economic-ecological model. This link is established through a social welfare function that incorporates the environmental damage (which is subjectively evaluated) along with the utility from consumption, the profits from production (pollution is positively correlated with production) and the costs from abatement. The social welfare function is defined either at the national or global level depending on the questions addressed.

The welfare function can be formulated with respect to either the emissions or the abatement of emissions. When emissions are used, the welfare function equals the difference between benefits from and damage to the environment, that is,

$$W_i = B_i - D_i = B_i\left(x_1, x_2, \ldots, x_i, \ldots, x_n\right) - D_i\left(x_1, \ldots, x_i, \ldots, x_n\right) \quad , \qquad (1)$$

where x_i denotes emissions associated with production in country *i*, with $i=1,\ldots,n$, B_i denotes country *i*'s benefits from the use of the environment for production and consumption activities at the international level, and D_i is country *i*'s damage resulting from pollution in all countries. When instead the variable is abatement, the welfare function takes the form

$$W_i = B_i - C_i = B_i\left(q_1, q_2, \ldots, q_i, \ldots, q_n\right) - C_i\left(q_1, \ldots, q_i, \ldots, q_n\right) \quad , \qquad (2)$$

where, q_i denotes abatement efforts in country *i*, and C_i is the associated cost of abatement.

The benefit function in (1) is analogous to the cost function in (2) and the benefit function in (2) to the damage function in (1). Thus, it is necessary that the term "benefit function" is always used in reference to the appropriate variable (emissions or abatement).

A number of assumptions are made in the literature regarding the welfare function in order to simplify the analysis while maintaining the important features of the specific problem each paper studies. We refer to these assumptions as model assumptions, in order to distinguish them from the assumptions associated with the form of the game and which will be called game assumptions.

The most important assumption concerns the form of interdependencies among countries. The benefits of country *i* are primarily associated with domestic pollution, that is in most cases it is assumed that $B_i\left(x_i\right)$. However, since we are concerned with international environmental problems, the damage each country suffers from pollution, or equivalently the benefit from abatement, is a function of all countries' emissions.

While each country's environmental damage unambiguously depends on all countries' emission, the same is not true for marginal damage. The degree to which other countries' emissions affect country *i*'s marginal environmental damage is very important in determining the level of cooperation among countries. A number of papers assume separability of the damage functions. This assumption implies that every country's marginal damage depends only on its own emissions and is independent of the emissions of all other countries, that is, the damage functions can be expressed as $D_i = \sum_{1}^{n} f_i(x_i)$. There are some works however, that assume that countries' marginal damage functions depend on other countries' actions. As it will be discussed latter, the relaxation of the independence assumption is particularly important in defining the outcome of the game, and thus it affects the level of international cooperation..

Another commonly used assumption is the symmetric country assumption. Although not realistic, assuming a common welfare function for all countries simplifies the exposition greatly.

A number of works concentrate on cost-effective analysis ignoring the environmental damage. In this case the problem reduces to cost-minimization subject to an abatement constraint. This approach has the advantage of avoiding the specification of the damage function, which is highly subjective and also difficult to use because of the non-linearities it exhibits. The drawback of this approach is precisely the lack of a damage function, which can reveal the incentives for cooperation as Barrett (1992a) points out. In the present paper we concentrate on works that employ welfare maximizing analysis.

The social welfare functions in (1) and (2) are static and are concerned with the flow of emissions. However, the majority of environmental damages result from the accumulation (stock) rather than the flow of emissions. The very important international environmental problem of climatic change, caused by the anthropogenic emissions of the greenhouse gases (carbon dioxide, chlorofluorocarbons, methane, nitrous oxides, and ozone) is the primary example of a stock externality. It is not the emissions of the greenhouse gases that cause the environmental damage, but rather the accumulated stock of these gases in the atmosphere.

In order to incorporate the notion of stock externality into the model, time is explicitly introduced. Environmental pollution becomes a dynamic process of accumulated emissions that dissolve slowly by natural processes (environment's assimilating capacity) which can be supplemented by abatement effort. Country *i*'s emissions $x_i(t)$ increase the stock of pollution over all countries $S(t)$ at time t. The concentration of pollutants damages the environment, and $D_i(S(t))$ describes country *i*'s damage. Country *i*'s benefits from production and consumption activities associated with the pollutant is described by $B_i(x_1(t),\ldots,x_i(t),\ldots,x_n(t))$. The welfare function under stock pollution becomes,

$$W_i = \int_0^{+\infty} e^{-rt}\left[B_i(x_1(t),\ldots,x_i(t),\ldots,x_n(t)) - D_i(S(t))\right]dt, \quad (3)$$

where *r* is the social discount rate.

The policy maker in country *i* is choosing the optimal path of emissions so as to maximize social welfare subject to the change of the accumulation of emission over time which is determined by the equation

$$\dot{S}(t) \equiv dS(t)/dt = \sum_i x_i(t) - d\,S(t),\ S(0) = S_0 \geq 0, \quad (4)$$

where, S_0 is the initial pollution level, $x_i(t)$ denotes country *i*'s emissions per unit of time and d reflects the constant rate of pollution decay by natural processes.

2.2 The nature of the game

Due to the apparent interdependence in countries' actions, in making its choices, each country takes into account other countries' actions. Countries' decision process can be modeled as either a cooperative or a non-cooperative game. When cooperation among countries is assumed, the focus is on the characteristic function of the coalition, that is, the objective is to maximize the total net benefits the coalition can share. On the contrary, when it is assumed that countries behave non-cooperatively, the objective is the maximization of each individual country's welfare. The latter is arguably a more realistic assumption and for that reason it is the one most commonly used.

Non-cooperative games can be formulated as (i) static games, (ii) repeated and (iii) dynamic games. Although only a small part of the literature is devoted to dynamic games, it is commonly accepted that dynamic games present a more realistic scenario. Furthermore, both repeated and dynamic games allow for the use of threats that may stabilize an agreement.

i. Static games (one shot games)

The most commonly used framework is a two stage game. In the first stage, the coalition stage, every country decides independently whether to join the coalition or not. In the second stage, the emissions stage, each country chooses its emissions level. The game is solved backwards, that is, the emissions stage is solved first followed by the coalition stage. Each country decides to join the coalition or not, fully anticipating the outcome of the related emissions game. (see for example Carraro and Siniscalco (1993)).

The main body of the literature examining the formation of international environmental agreements within a two stage game framework uses the following set of assumptions (as mentioned in Carraro and Siniscalco (1998)).

1. Decisions are simultaneous in both stages.
2. Countries are presented with a single agreement.
3. When defecting from a coalition, a country assumes that all other countries remain to the coalition.
4. The payoff function of each country increases monotonically with respect to the size of the coalition.

The first assumption excludes the possibility that one of the countries has any advantage in the game. The second assumption excludes the formation of more than one coalition. Relaxing this assumption leads to the formation of more than one coalition even in the case of symmetric countries (see Carraro and Siniscalco (1998)). The third assumption, consistent with Nash conjectures, excludes the possibility of farsightedness. Relaxing this assumption yields a different equilibrium, which we discuss briefly below. Finally, the fourth assumption is a model and not a game assumption. If we allow the payoff to be decreasing in the number of signatories after a number of countries sign the agreement, the optimal number of signatories may be less than the maximum stable coalition number.

Alternatively, when the expansion of coalitions is examined the game can be reduced to a one-stage game using a partition function that replaces the emission game (see for example Carraro and Moriconi (1997)). The partition function assigns to each coalition structure a vector representing the worth of all coalitions within this structure. Thus, countries decide whether to join an already existing coalition knowing all possible payoffs. In this case, in addition to the above-mentioned assumptions, three more assumptions are required.

5. The Nash emissions game has a unique equilibrium for every coalition structure.
6. Within the coalition players play cooperatively while the coalition and single countries compete in a non cooperative way among them.

Assumptions 5 and 6 are crucial in defining the partition function, while another assumption is also used to simplify derivation of the partition function by introducing an equal sharing rule among coalition members.

7. All countries are identical, that is, they have the same strategy space in the emission game (Carraro and Moriconi (1997)).

Games with coalitional equilibrium

Although most of the literature develops around the above mentioned assumptions, there is a set of papers that work within a considerably different framework. The basic three assumptions are dropped. Countries do not decide simultaneously, allowing for deviations until the (permanent) equilibrium is reached. However, this is not formulated as a repeated game since countries realize payoffs only at the equilibrium. Countries are not presented with a single agreement, since many different coalitions can emerge (this is a more realistic assumption). And finally, when a country or a coalition deviates, it does not assume that all other countries or coalitions will retain the same policy.

Within this framework Ecchia and Mariotti (1998) examine self-enforcing International Environmental Agreements based on the notion of coalitional equilibrium (see also Mariotti (1997)). For any coalition S, a coalitional strategy for S specifies the proposal of coalition S at any possible status quo. A coalitional equilibrium has the following properties: 1. For an equilibrium strategy profile, no coalition wants to deviate from it. 2. No coalition should be able to gain by switching to an alternative coalitional strategy, at any status quo (Nash and sub-game perfect equilibrium requirement). 3. Deviations from the current status quo should be motivated, and 4. At the equilibrium point, no player can be forced to a worst payoff

When a coalitional equilibrium is reached, then we have a self-enforcing IEA. This framework captures the notion of "farsightedness" which might be a more realistic environment. Carraro and Siniscalco (1998) and Carraro and Moriconi(1997) consider this kind of changes and analyze the consequences of principals like "coalition unanimity" and "rational conjectures" within the framework of one shot non-cooperative games. The assumption of rational conjectures is very close to the notion of farsightedness.

ii. Repeated games

In reality the decision to join or leave an agreement can be reassessed at different points in time, in light of new facts and the knowledge of other players' past behavior. Assumptions 1 and 3 are relaxed, while assumptions 2 and 4 are usually kept. The main advantage of repeated games is that they can incorporate credible threats, and thus, increase greatly the potential for a cooperative solution. The treats are formulated in terms of abatement, that is, the observant countries threaten to return to the non-cooperation stage when one country is detected to cheat on the agreement. The treat could take various forms. It could be a trigger strategy turning to non-cooperation forever after one country cheats. The severity of the trigger strategy though often renders it non credible. Or it could be a tit-for-tat strategy that implies non cooperation only for one period. Because the credibility of punishments under the tit-for-tat strategy, these strategies –or modifications of these strategies— have been more often considered in the literature (see for example Barrett (1994) and Heal (1994)).

Dynamic games

Although repeated games are doser to reality, they assume that the parameters of the game remain the same over time. However, pollutants accumulate over time, and thus, the damage to the environment is changing over time. Technology and other parameters of cost and damage functions are changing over time too. Thus, the most appropriate framework to captures these facts is a dynamic game framework. It is based on the dynamic welfare function presented above and assumes a repeated negotiation process.

3. Basic notions

This section discusses the basic notions of the game-theoretic approach to analyzing international environmental agreements. The outcomes of different strategies available to countries are compared to the following benchmark cases. At the one end lies the full cooperation equilibrium while a the other end there is the business as usual scenario or the Nash equilibrium. The most commonly used definition of self-enforcing agreements is also presented.

3.1 The benchmarks

Full cooperation

There are some cases, which can serve as reference points when we discuss the possible equilibria for international environmental agreements. At the upper end lies the full-cooperation solution requiring the maximization of aggregate welfare, presented as the sum of all counties' welfare. This is the efficient solution, at which, marginal cost across countries are set equal to the aggregate marginal damage, that is, $\sum_{i=1}^{n} MD_i = MAC_i, \forall i$, where MD_i and MAC_i denote marginal damage from emissions and marginal abatement cost for country *i* respectively.

This solution could be reached in the presence of an international authority that has the necessary information to derive and the power to enforce the efficient allocation of abatement efforts. However, no such authority exists at present, and it is not realistic to assume that countries would relegate such power to an international body in the future. Alternatively, countries could voluntarily negotiate for and participate to a cooperative agreement. However, there are two main problems making the achievement of full cooperation very difficult if not impossible. The first is that, at the full cooperation equilibrium some countries may be worse off compared to the non-cooperation case. This problem could be sidestepped with the use of the appropriate transfers to countries that are loosing from cooperation. Since total gains from full-cooperation exceed aggregate losses, countries that realize the gains could certainly compensate the loosing countries. The second problem is that even if all countries gain from full cooperation, they have an incentive to free ride on the agreement. A county that decides to cheat by decreasing its abatement efforts lowers its cost while realizing the benefit of lower aggregate emissions achieved by the rest of the participants to the agreement. Although all countries could be better off at the full-cooperation solution, none will choose to cooperate given the existing incentives, at least within the static framework.

Although not realistic, the full cooperation solution is used often as a benchmark case, depicting the first-best world, against which other possible solutions are compared. Alternatively, the Pareto-dominant solution is used as the benchmark case. This solution is derived by maximizing aggregate welfare subject to the condition that no country is worse off when it cooperates. Although it presents a more realistic scenario, it is also subject to free riding problems and furthermore, it does not lead to first best since it is a constrained problem.

Best reply functions

Two types of free riding behavior could be considered depending on the assumptions regarding countries' best reply functions. The best reply functions present one country's optimal level of emissions as a function of the other countries' emissions. They reflect one country's marginal emission damage, as determined by other countries' emissions. The best reply functions are orthogonal when the emissions of a country do not affect the marginal benefits of other countries (independence assumption). In this case, countries outside the coalition benefit from the reduction in emissions achieved by the cooperating countries but they cannot affect the benefits derived by the members of the coalition. Conversely, when interdependency between countries' emission strategies is assumed, the best reply functions are negatively sloped. That is, a country will reduce its emissions in response to an increase in other countries' emissions, and this reduction will be greater the higher is the elasticity of its best reply function and vice versa. In this case, free riding countries not only enjoy the benefits from the coalition but they can also hurt the stability of the coalition. This is because a decrease of emissions by the cooperating countries will be offset by an increase of emissions by the non-cooperating countries. In such case some members of the coalition may lose and thus, decide to leave the coalition. Therefore, the assumption regarding the best reply functions is crucial for the stability of partial agreements. The impact of different degrees of emission leakage, that is interdependencies in countries' emission strategies, on the form and the size of the coalition is examined as well as proposals for limiting the leakage effects.

Nash equilibrium

At the lower end of possible equilibria for international environmental agreements lies the business as usual solution at which countries do not engage in abatement activities. This solution would be reached assuming that countries are ignorant of the damage caused by their own emissions (see Fankhauser and Kverndokk (1992)). When, however, countries realize this damage they have an incentive to reduce their emissions to some degree even in the absence of other countries' action. Thus, the business as usual solution is far too extreme to even be used as a benchmark. The lower end could instead be depicted by the Nash equilibrium, which is derived assuming that every county chooses its abatement effort given the other countries' choices. At the equilibrium each country sets its marginal abatement costs equal to its marginal pollution damage, that is, $MD_i = MC_i, \forall i$. It is clear that the Nash equilibrium does not achieve either efficiency or cost minimization. The Nash equilibrium is sometimes referred to as a "threat point", that is, as the default situation in case no agreement can be reached (see for example Folmer et al (1998)).

3.2 Self-enforcing International Environmental Agreements

Because of free riding problems, neither the full cooperation nor the Pareto optimum solution can be considered equilibrium outcome. The literature has focused on identifying a sustainable equilibrium that will present an improvement over the Nash equilibrium. Such a solution must be self-enforcing since external enforcement is impossible. The vehicle of the analysis is non-cooperative game theory. The most important result shared by the main body of this literature is that coalitions are formed including only a fraction of the countries. The exact number of the countries in the coalition and the level of the abatement are determined by the particular assumptions of each model.

In order for an international environmental agreement to be stable, there should be some mechanism rewarding entry and penalizing exit from the coalition. Such schemes have been proposed in Carraro and Siniscalco (1993) and Barrett (1994). According to them when a country joins the agreement the other signatories increase their abatement levels, rewarding the joining country. When a country defects, the opposite happens and the defecting country is punished. These punishments are credible because they result from the signatories' welfare maximization problem. The problem is that the magnitude of the punishments is not always big enough to sustain an agreement with many countries.

Carraro and Siniscalco (1993) present the following two conditions for achieving a self-enforcing agreement. The profitability condition requires that profits of every country belonging to coalition *s* exceed its profits outside the coalition, that is $P_i(s) \succ P_i^0$, $\forall i \in s$, where $P_i(s)$ are country *i*'s profits, when the coalition *s* is formed, and P_i^0 are country *i*'s profits when the coalition has no members. There are two stability conditions. The internal stability requiring that no country wants to defect, that is, $Q_i(s-1) < P_i(s)$, where $Q_i(s-1)$ are country *i*'s profits when it defects from coalition *s*. The external stability requiring that no other country wants to join the coalition, that is, $P_j(s+j) < Q_j(s)$, $\forall j$ that does not belong to coalition *s*, where $P_j(s+j)$ and $Q_j(s)$ are country *j*'s profits when it joins in and when it remains outside coalition *s* respectively.

The intuition of the profitability condition is that no country will join the coalition if it does not gain, compared to the non-coalition case. This condition is fulfilled most of the times except when there are leakage effects, or great asymmetries across countries. The stability conditions are imposed to eliminate the free-riding incentives which are present even if a country gains from a coalition, since it may gain more by defecting, that is, by free riding on others' abatement efforts. The internal stability condition ensures that if a country defects, the profits from free riding will be less than the profits if it stays in the coalition. This is achieved by reducing the non-defecting countries' abatement level as a response to defection. Since this response results from the countries' maximization problem it presents a credible threat.

4. The main issues examined

We present the main developments in the literature that examines the formation of international environmental agreements within a game theoretical framework by focusing on three main issues. The first is the size of stable coalitions, the second is the possibility of expanding stable coalitions and the means by which an expansion can be achieved, and the last concerns the choice of the policy instrument to be used by policy makers.

4.1 Determining the size of a stable coalition

Defining the number of signatories in international environmental agreements is the most important task. Can there be a stable full coalition, the so-called "grand coalition", or are the only stable international environmental agreements those with a small number of signatories? One strand of the literature based on cooperative game theoretic analysis shows that the "grand stable coalition" exists (see Chander and Tulkens (1995), (1997), Tulkens (1998) and Uzawa (1997)). Another strand of the literature shows that, within a non-cooperative game theoretic framework, only small number coalitions are stable (see Barrett (1994), Carraro and Siniscalco (1993)). Finally, support for the existence of a stable full coalition is provided within the framework of non-cooperative game theory, allowing for repeated games and assuming that countries are farsighted enough to understand the full consequences of free-riding behavior (see Ecchia and Mariotti (1998a and b)).

Chander and Tulkens show that maximization of the characteristic function, that is, the aggregate welfare of all countries --given in either equation (1) or (2)--, can be achieved using an optimal set of transfers. The optimal transfer is written as function of country *i*'s emissions at the optimum x_i^* and the Nash equilibrium $\bar{x}_i$, that is, $T_i = -\left[g_i\left(x_i^*\right) - g_i\left(\bar{x}_i\right)\right] + MWPE_i / MWPE_n \left[\sum_{i=1}^{n} g_i\left(x_i^*\right) - \sum_{i=1}^{n} g_i\left(\bar{x}_i\right)\right]$, where $g_i\left(x_i\right)$ denotes production in country *i* as function of emissions, and *MWPE* stands for the marginal willingness to pay for the environmental good (see Tulkens(1998)). Country *i* is compensated for any loss from participating in the grand coalition, and it takes in a proportion $MWPE_i / MPWPE_n$ of the total benefits achieved by the grand coalition. It is shown that once the grand coalition supported by the transfer scheme is reached, it cannot be improved upon by any sub-coalition, that is, the grand coalition is both profitable and stable. The threat that sustains the grand coalition is the "Nash behavior" of all the remaining countries against the defecting countries and this is an individually rational threat (see Tulkens(1998)). One issue that is not addressed though is the subgame perfection of this threat, which is crucial for its credibility. Another critical assumption of the above analysis is that countries are first presented with the option of the grand coalition, that is, there is no discussion of the process by which the grand coalition is formed. It appears that the existence of an international institution that designs and proposes the transfer scheme to individual countries is assumed.

Within a different framework and contrary to the above result, Barrett (1994) shows that only small number coalitions are stable. His analysis is based on the notion of self-enforcing agreements using non-cooperative game theory. The following assumptions are employed: countries are identical (symmetry assumption); cost functions are independent of other countries' emissions (independence assumption); there is full information regarding benefit functions; and abatement levels are observable. It is further assumed that signatories maximize their collective welfare, while the non-signatories behave individually, taking all other players' actions as given. In defining the number of signatories at the self-enforcing equilibrium, the two stability conditions discussed above are imposed. The self-enforcing international environmental agreements are examined both within static and infinitely repeated game framework.

Within the static framework, and using simulations on particular cost and benefit functions, Barrett shows that the number of signatories depends on the ratio of the slopes of marginal cost and benefit functions. When this ratio is large, the self-enforcing coalition has only few members and vise versa. However, when the number of signatories is large, the total benefits from cooperation are small. Thus, when the benefits from cooperation are potentially large the number of signatories of a self-enforcing coalition is small, and thus, the potential benefits are never realized. The result that only few countries will be joining a stable coalition is quite robust since it is reaffirmed in a variety of specifications by Hoel (1991), Carraro and Siniscalco (1993) and Botteon and Carraro (1998). Similar results are derived when the assumption of symmetric countries is dropped. Barrett (1993) shows that even in this case, when stable coalitions are possible the benefits are small and vise versa. Botteon and Carraro (1998) also show that only a small number of countries form a stable coalition in the absence of transfers and that these countries are the ones that suffer the higher environmental damage.

What is more surprising is that similar results derive even when international environmental agreements are examined as repeated games. Allowing for infinitely repeated games and using a small interest rate one would expect that the increase of credible penalties would lead to increases in the number of signatories. However, following Farrell and Maskin (1989) in defining renegotiation-proof agreements and using specific cost and benefit functions, Barrett proves that there is an upper limit to the number of countries that can sustain the full cooperative outcome. This limit is high only when the difference in global net benefits between the non-cooperative and the full cooperative outcomes is small. When this difference is large only few countries can sustain the full cooperative outcome.

These results are similar to those derived within the static framework, verifying the disappointing result. Although the particular number of the signatories of a self-enforcing agreement depends on the functional form of the welfare function, the main result remains the same. Barrett sites the Montreal Protocol as an example supporting his findings. Although the participation in the Montreal Protocol is large (around 140 countries), the net benefits to individual countries are small.

Finus and Rundshagen (1998) examine the stability of the grand coalition in relation to two different instrumental choices, the uniform emission reduction quota and the uniform effluent tax. While we defer the discussion of the comparison of the two instruments for section 4.3., we present their findings regarding the stability of international environmental agreements. They assume asymmetric countries and use cost-benefit analysis first within a static and then within a dynamic framework. They assume that negotiations have the Nash equilibrium as a starting point and that the Lowest Common Denominator rule is used. This means that the prevailing proposal is the lowest abatement proposal, making sure that every country signs the agreement. These proposals are the outcome of welfare maximization for every country. They find that regardless of the policy instrument used the grand coalition is unstable. They also test whether the stability problem can be solved if the game is repeated over time. Using an infinitely repeated game framework and the weakly renegotiation-proof conditions proposed by Farell and Maskin (1989), they arrive at results similar to those derived by Barrett (1994), that is, whenever the potential gains from cooperation are large, there is a stability problem. The social optimum solution is only stable in a few cases (when the number of countries is small and the benefit-cost ratio is large), for which abatement targets are low. This result holds for small discount rates too. Hoel (1992) also agrees with this result showing, using numerical simulations, that very few countries will form a stable coalition. Furthermore, Finus and Rundshagen find that sub-coalitions might be more efficient than the grand coalition because of the asymmetries that countries have.

Ploeg and de Zeeuw (1991) present the problem of international cooperation as a differential game, that is, countries choose their strategies at any point in time depending on the concentration of pollutants at that time. They use a differential model similar to the one described by equations (3) and (4) and they assume that countries are of equal size sharing the same fixed emission output ratio and the same damage function. They compare this case to the case that countries commit to a strategy at the beginning of the game. In the latter case, using optimal control techniques they show that the stock of pollutants is bigger under the non-cooperative open loop Nash equilibrium compared to the full cooperation outcome. Next they argue that it is more realistic to assume that countries do not decide once and for all, but change their decisions according to the current observation of the stock of pollutants. Under this assumption we have to use the feedback Nash or sub-game perfect Markov equilibrium concept. Applying this concept and using Bellman's dynamic programming they find that the concentration of pollutants is higher relative to the open loop Nash outcome.

The intuition is that, when a country observes an increase in the concentration of pollutants it assumes that the other countries will abate more and so underestimates the marginal damage to the environment and thus pollutes more. Ploeg and de Zeeuw conclude that the unrealistic assumption that countries commit to abatement strategies at the beginning of the game, leads to substantial underestimation of the damage to the environment and thus, to the underestimation of the benefits from cooperation.

The existence of a large number of international environmental agreements and the great number of signatories in many of those, such as the Montreal Protocol, is the main motivation of the work by Jeppesen and Andersen (1998). To explain the large number of participants in international environmental agreements they extent the model developed in Barrett (1994) to incorporate the idea of fairness. In so doing, they utilize Rabin's (1993) analysis of embodying non-material aspects into players' payoffs. In particular, they add to the total benefits derived by a non-participating country a loss function from not joining the coalition. The loss from not joining the agreement becomes larger the higher is the number of participating countries. Within this framework, the size of the coalition depends on the particular functional form of the loss function. It is shown that for some loss functions, previously stable coalitions collapse, while for other functional forms even full cooperation is possible. Thus, depending on the weight countries' assign to the non-material loss of not participating in the coalition the stable coalition could be either small or large.

Crucial role for the existence of a partial coalition plays the slope of countries' best reply functions reflecting the assumptions regarding the interdependencies among countries' actions. If best reply functions are negatively sloped, then any decrease of emissions from the cooperating countries will be offset by an increase from those outside the agreement. These are the so called "leakage effects" and reduce the strength of partial coalition. If on the other hand, best reply functions are orthogonal then the non-cooperating countries simply enjoy a better environment without offsetting the other countries' effort.

Heal (1994) shows that if the independence assumption is dropped or there are fixed abatement costs then countries could unlock from the prisoner's dilemma situation, because there are increased incentives to cooperate. The analysis introduces the notion of minimum critical coalition to help define coalition stability. The minimum critical coalition without side payments is defined as the coalition in which the benefits to individual members are at least equal to costs. This is very similar to notion of profitable coalition discussed above. Allowing for side payments, countries with net benefits support net losers in order to widen the size of coalition.

Using the notion of minimum coalition and a simplifying two-country model, Heal shows that countries have greater incentives to join and remain in a coalition when the independence assumption is relaxed, that is, when costs and benefits of abatement in one country depend on other countries' actions. Heal argues that this assumption is more realistic since first, an increase in one country's abatement affects the marginal net benefit from abatement in other countries and second, the marginal cost of abatement in one country decreases with the abatement activities of the other countries. These reinforcement effects lead to higher level of abatement and limit the problem of free riding. Similar results derive in the presence of fixed abatement costs.

The leakage effects and finding ways to suspend them is also the concern of Hoel (1994). He assumes that only a small number of countries participate in an international environmental agreement and thus, only a sub-group of countries is committed to cooperation. Using a general equilibrium model he examines ways to limit the leakage problem within a non-cooperative static game framework. He shows that in most cases a mix of demand and supply policies gives the better outcome. He also shows that the existence of transfers is in the interest of both cooperating and non-cooperating countries, through which the cooperating countries affect the consumption and production of carbon in the non-cooperating countries.

Ecchia and Mariotti (1998a,b) motivated by the contradiction between the theoretic results showing that only small number agreements are self enforcing and the reality of large number agreements, extend the basic idea of self-enforcing international environmental agreements. They identify two problems with the basic model of self-enforcing agreements. In the basic model, countries are assumed to behave myopically, that is they ignore other countries' reaction when making their choices. If instead countries are farsighted, that is they can anticipate other countries' reaction to their choices and incorporate them into their decisions, the logic of the two stability conditions is undermined. The second problem is that the basic model does not allow other coalitions to form. Allowing countries outside the coalition to act as groups with respect to the main coalition again undermines the use of the stability conditions.

Ecchia and Mariotti redefine self-enforcing international environmental agreements incorporating these two elements into a static model that captures the necessary recursiveness of farsightedness. The notion of the "coalitional strategy" is introduced, upon which rests the definition of "coalitional equilibrium" which is a kind of subgame perfect equilibrium for coalitional strategies. Farsightedness is built into the model since any coalition is considering the full consequences of its moves by deriving expectations regarding the reaction of the other countries or coalitions. Using the simplifying case where only three countries are involved in the game, they identify four different versions of the game. In the strong and weak prisoner's dilemma cases polluting is the dominant strategy, but in the weak version there is some incentive to cooperate within sub-coalitions. In the strong and weak chicken game there is stronger incentive to cooperate relative to the prisoner's dilemma cases.

The analysis of these four types of games shows that complete non-cooperation is impossible, while full coalition is always a possibility. The intuition of this striking result --especially for the strong prisoner's dilemma case-- is that when farsightedness allows countries to recognize the credibility of the free-riding threats they choose to cooperate. It is the "balance of credible threats" that keeps the full coalition alive in all four cases. Thus, when the idea of farsightedness is injected into the model, the possibility of larger coalitions increases.

Endres and Finus (1998) examine the stability of international environmental agreements by incorporating in the welfare function the environmental awareness as a factor of the total costs of emissions. Within the static game framework they show that total gains from cooperation rise with the level of each country's awareness. On the contrary, the stability of the grand coalition decreases, since the gap between Nash equilibrium and the social optimum outcome increases for the countries with lower awareness. The latter result stands even for a small proportional increase in countries' awareness. The stability of the coalition increases only for a "big enough" rise in environmental awareness for each of the participants.

Next, they check for the validity of these results in a dynamic context. They use the infinite supergame framework and in order to simplify the analysis they assume symmetric countries with respect to environmental awareness. They restrict their attention only to subgame perfect strategies. They again find that when the awareness of one country increases, the instability of the grand coalition increases as well. Contrary to the results within the static framework, they show that a small proportional rise in each country's awareness increases the stability of the coalition while the reverse holds for a "big enough" increase. The explanation is that a big increase in awareness decreases again the gap between Nash equilibrium and the social optimum outcome and so punishment for the defector (playing Nash) is less severe. Taking the example of an agreement based on a uniform emission reduction quota they find similar results. The increase of environmental awareness in one country affects both the incentives to participate in an agreement and the punishment for defection for the other countries. Those effects in general work in opposite directions and the “punishment effect” dominates the “incentive” effect, in a way that the final outcome of an increase in awareness can be the increase of the instability of the agreement.

Endres and Finus examine also the effects of unilateral abatement actions in the case of quota agreement. They distinguish four cases. In the first, the country takes the unilateral action after the agreement has been signed and the demand for unilateral action by the country’s citizens is raised after the signing of the agreement. In this case the country that takes the unilateral action looses but there is a total gain. The three other possible cases are examined are: unilateral action after the agreement is signed but the demand was raised before; unilateral action before the signing of the agreement and when the country does not take unilateral action but the demand is raised before the signing of the agreement. In all the three cases the country examined suffers a welfare loss while there is a total loss too. This is because the rest of the countries sign a less demanding agreement knowing or anticipating (rightly or wrongly) the country’s unilateral action. Thus, Endres and Finus show that good intentions are not always enough and that they might even worsen the situation.

Na and Shin (1998) introduce uncertainty about the environmental benefits and examine its effect on coalition formation. They assume that each country’s marginal benefit of aggregate abatement is a random variable with a known distribution. They use a three country model with separable benefit and cost functions and they examine the outcome of negotiations before nature reveals the real value of marginal benefit for every country (ex ante negotiations) and after this value is revealed (ex post).

They show that ex ante negotiation lead to the formation of the stable grand coalition (always referring to their three country model). On the opposite, ex post negotiations do not lead to a stable grand coalition. Depending on the distribution of marginal benefit they can lead either to a partial coalition or to no coalition at all. The intuition for the superiority of ex ante negotiations is that when countries don’t know everything, they are closer to each other, but when they know what exactly they have to gain and loose, they have different interests that makes cooperation difficult.

Finally they mention a second advantage of ex ante negotiations, that is, time savings. Countries do not have to wait until nature reveals the value of marginal benefit and this can prevent irreversible damages to the environment.

4.2 Expanding stable coalitions

The main body of the literature reviewed above shows that only small coalitions are stable except if the model is extended to include other issues. The use of abatement reductions or increases to punish defecting countries or attract new countries has limited effect. Thus, the next step is to examine other measures for expanding the initial stable coalition. The measures that have been considered in the literature are used to either increase the participation incentives through transfers and linkages to other issues or decrease the gains from cheating on the agreement through penalties.

Transfers and Commitment

As discussed in the previous section, transfers can be used to increase participation in international environmental agreements. That discussion though was based on cooperative game theory that implicitly assumes the existence of a supranational authority that enforces the transfer scheme. In this section we review some other ways of modeling transfers and then we discuss ways in which countries or group of countries could make credible commitments.

Chander and Tulkens (1992) propose a different way of modeling negotiations. Instead of directly computing the preferred optimum (which requires the complete knowledge of benefit and cost functions) and finding ways to achieve it, they propose a “gradual” procedure. By this procedure, at each stage of negotiations, only gradual improvements of the previous situation are to be decided, which are Pareto-improving and converge to a social optimum. In this way only the local properties of the above functions need to be known. They model this procedure specifying a system of differential equations, which provide in every round of negotiations the abatement level for each country for the next step. The only required data are counties’ marginal willingness to pay for environmental improvement and marginal cost of abatement, at each point in time. The repetition of this process can lead close to an optimum.

In their model transfers are allowed. As they point out, they are not necessary for the optimality of the final outcome but for supporting individual rationality. They propose a distribution rule that is first proposed by Chander and Tulkens (1991). According to this rule each country is burdened by a proportion of total abatement cost. This proportion equals the country's intensity of preferences for environmental quality relative to the intensity of total preferences.

A weak point of this analysis, which is based in cooperative game theory, is the free-rider behavior that might occur by some countries and for which no solution is given inside this framework so far, as the authors mention.

Botteon and Carraro (1998) examine the possibility of expanding an initially small coalition in the case of non-identical countries, their asymmetry defined on the basis of differences in their marginal environmental damage. They focus on comparing the Nash bargaining rule with the Shapley rule on the basis of their ability to achieve the expansion of the coalition. Using simulations of a five country model, they show that self-financed transfers can expand the initial stable coalition under both burden sharing rules if some degree of commitment is present. Furthermore, they show that under the Shapley rule expansion is possible even without partial commitment. This is because the Shapley rule distributes the gains more evenly, offering higher stability to any coalition. Using simulations they find that a coalition with transfers based on Shapley values Pareto-dominates a coalition based on Nash bargaining solution.

Using cost-benefit analysis within a framework similar to the one used by Carraro and Siniscalco (1993), Hoel and Schneider (1997) examine the effect that a system of side payments has on the countries' incentives to cooperate. They assume symmetric countries (same production function and equal marginal environmental costs for every country) and they incorporate non-environmental costs for not signing an agreement, which differ from country to country. They find that in the absence of side payments and with the help of country specific non-environmental cost function the number of signatories is not very small. In the specific example they use the number of signatories is half of all countries. On a contrary, when side payments are used, fewer countries join the agreement and the influence of non-environmental costs is diminished.

The existence of a large number of international environmental agreements and the great number of signatories in many of those, such as the Montreal Protocol, is the main motivation of the work by Jeppesen and Andersen (1998). Employing the exact same model developed in Barrett (1994) they identify two possible directions in which this model can be extended to allow for higher participation in international environmental agreements. First, they assume that commitments by individual countries or group of countries are credible, contrary to Barrett who strongly opposes the idea that such commitments can be credible. Allowing a group of countries to make credible commitments regarding their abatement implies that this group of countries assumes a leader's role in forming the coalition. This leading role allows them to calculate potential aggregate benefits from enhancing the coalition, and devise side payments to countries that have a follower's role in order to achieve the optimum participation. Simulating Barrett's model under the assumption of credible commitments Jeppesen and Andersen show that a relatively small core of committed countries could expand the coalition substantially, achieving even full coalition. Although very interesting, the idea of commitment suffers from the absence of any explanation of how the power of commitment develops.

After proving that only a fraction of countries would cooperate in the first place forming a stable coalition, Carraro and Siniscalco (1993) search for ways of expanding a stable coalition. They assume complete information, symmetric countries and that agreements are non-binding and use the profitability and stability conditions to define the stable coalition. The game has the usual two stage form, that is, countries choose first whether to cooperate or not and subsequently they choose their emission level. Within this framework, they show that when there is no commitment to the cooperative strategy from the countries belonging to the stable coalition, self-financed transfers cannot expand the initial coalition. Then, they examine four types of commitment that could lead to the expansion of the core coalition. The types of commitment examined are: stable coalition commitment (only the core countries commit), sequential commitment (every new member commits), full-cooperation minimum commitment (committing countries are enough to induce full cooperation), and external commitment (commitment from a subset of countries outside the core). They show that whenever any of these types of commitment exist, the cooperation decision does not lead to a prisoner's dilemma situation. The actual number of countries that can be attracted to the core coalition depends on the type of commitment (for example, sequential commitment produces wider coalitions than stable coalition commitment) and on the slope of the best reply functions (the more negative the higher the incentive to cheat the coalition).

Petrakis and Xepapadeas (1996) extend the argument of partial commitment developed by Carraro and Siniscalco (1993) to the case of non-identical countries and asymmetric information. Assuming linear damage functions, separable for every country (independence assumption), and the existence of a stable coalition able to commit to cooperation, they show that expansion of the initial coalition is possible with self-financed transfers. Countries with the higher marginal environmental damage form the initial coalition and try to expand it using welfare transfers to countries suffering lower damages. The stable expansion of the coalition is achieved if the number of the member to the initial coalition is large and/or their benefits are high relative to the disutility of the new members. The stability of the expanded coalition depends on the ability of the initial group of countries to monitor and enforce the agreement to the new members. In order to avoid the high costs associated with monitoring and enforcement, Petrakis and Xepapadeas propose a mechanism for overcoming the moral hazard problem. This mechanism is based on inducing voluntary actions by the new members of the coalition that have a low cost of monitoring their emissions. This mechanism, which can be thought as an alternative to an international authority, is effective in resolving the moral hazard problem and it has low administrative cost.

Issue linkages

As the above discussion reveals transfer payments can substantially enhance the size of a coalition. When direct transfer payments are not possible, participation to an environmental agreement can be linked to other issues in such ways as to increase the benefits from participation. Such issues include trade policy, research and development, international debt and development assistance. Although there is some work on linking environmental agreements to international debt (see for example Mohr and Thomas (1998)) most of the literature focuses on trade and R&D linkages. These two types of issue linkages are reviewed in what follows.

Motivated by the fact that trade sanctions were the only enforcement mechanism for signing the Montreal protocol, Barrett (1997) uses a partial equilibrium model to examine the effectiveness of trade sanctions in signing an international environmental agreement. He considers trade only in goods that are linked to the environmental problem and models the game as a static one with four stages (similar to the two stages one). Using simulations, Barrett shows that the results for the environment are better when there are trade sanctions.

The linkage of trade policy with the environment is the subject of Breton and Soubeyran (1998). They use a detailed micro model where production and consumption is explicitly examined and where the policy maker chooses the technology standard --the only available policy instrument--, that balances production efficiency and environmental protection. They also deviate from previous works in considering two different policy rules: social optimization and median voter policy. Within the autarky framework they show that the technological standard is more demanding under the median voter policy. When trade is examined, they show that there is a strong relation between trade policy and the environment, whenever the tradable good affects the environment. Assuming that countries are large enough to influence the terms of trade, they examine the role of tariff wars in environmental negotiations. They show that trade policy can sustain environmental cooperation only if it is subgame perfect, that is, the "good" environmental country can win a tariff war if it is needed.

The expanding of small stable coalitions with the help of issue linkage is the purpose of Carraro and Siniscalco (1998) article. The linkage that is studied is through R&D cooperation. Their model incorporates not only countries' but also firms' behavior. Firms are assumed to maximize their profit, while countries the sum of domestic firms' profit, consumer surplus and environmental quality. Firms choose the level of their R&D expenditure and production given the demand and a cost function. Countries decide whether to join the international environmental agreement or not and subsequently the abatement level that will be imposed on the local firms. Thus, the game is played in three stages, the first two describing countries' choices and the last one the choice of R&D and production by firms.

It is assumed that the R&D spillovers for a cooperating country are bigger compared to the spillovers of a non-cooperating one, that is, R&D benefits is an excludable good. It is also assumed that firms from the cooperating countries have lower marginal costs than the other firms and that governments' best reply functions are orthogonal. Another important assumption is that technological progress affects both the economic and the environmental features of production.

They solve the game backwards and they use the usual profitability and stability conditions. Using simulations they show that linking the R&D with the environmental negotiation procedure leads to a great increase in the stability of the environmental cooperation. It can increase the number of signatories and the total welfare of the cooperating countries.

Katsoulacos (1998) is also studying the same issue linkage, but in a different way. He argues that R&D cooperation is not an issue dealt with at the country level.

Instead, it is firms that cooperate and governments cannot stop them. Governments could promote domestic firms' participation in research joint ventures using subsidies. Since markets cannot sustain the socially optimal level of R&D investments, countries could gain from subsidization of investments in research joint ventures. If a country does not gain from unilateral subsidization of these investments, then there is a space for joint efforts and an international agreement on R&D is both profitable and stable. If this agreement is linked with an environmental one, the latter becomes stable too in the case that the gains of defection from the environmental treaty are smaller compared to the gains from R&D cooperation.

The above discussion indicates that expansion of the initial coalition is possible in a variety of ways, some of which can produce very satisfactory results. We now turn to the question regarding the instruments of environmental policy.

4.3 Choice of policy instrument

The choice of the policy instrument for implementing an international environmental agreement is crucial. Environmental policy instruments are usually classified as either market based, which includes emission taxes and tradeable emission permits, or command and control, such as technology standards and firm-specific emission quotas. The economic literature favors the use of market based instruments because they can achieve both static and dynamic efficiency. Static efficiency refers to minimization of total abatement costs which is achieved through efficient allocation of abatement effort among sources. Dynamic efficiency refers to technological improvement over time, which is achieved through the provision of the necessary incentives. However, both static and dynamic efficiency arguments are undermined in a second best world, as for example in the presence of market imperfections.

Similar arguments carry to the discussion of international environmental problems. In the international context, it is crucial to consider the incentives that each policy instrument provides to countries for participating and complying with the agreement. Furthermore, considering linking environmental problems to other issues such as trade, R&D or international debt (Mohr (1995) examines the latter case using tradeable emission permits) the issue of choosing the policy instrument becomes a second best problem. In such cases careful design of the market based policy instruments is paramount.

Finus and Rundshagen (1998) examine two different policy instruments, namely the uniform emission reduction quota and the uniform effluent tax. They use the lowest common denominator rule briefly discussed in section 4.1. They show that total welfare is higher compared to the Nash equilibrium under both policies. However, under the uniform quota policy the benefit cost ratio of emissions is higher relative to the tax policy. The advantage of the quota policy is that it burdens less the low environmental conscious countries, making their participation easier, compared to the tax regime. Thus, although the uniform tax policy would yield higher total welfare because of its cost effectiveness, the uniform quota policy dominates due to the lowest common denominator rule. Within an infinitely repeated game framework the show that in a second best world, the quota policy is preferred to the tax policy. They also find that inside sub-coalitions the tax regime is preferred by the signatories only when their number is small. For a large number of countries with heterogeneous interests they prove that the quota regime is preferred, because it provides a more even distribution of abatement burdens.

Hoel (1992) uses the non-cooperative static game framework and assumes that countries have identical cost functions while their evaluation of abatement benefits differ, that is, they have different damage functions. Within a specific example he finds that when a uniform reduction in emissions is adopted, half of the countries that signed the constrained social optimum abatements won't sign now because the profitability of the agreement is severely reduced for those countries. This leads to a reduction in total welfare and to an increase in emissions from the cooperating countries. In the case that the cooperating countries will prefer the proposal of the median country, the situation is even worse and complete non-cooperation might be the final outcome. So he concludes that a uniform emission quota is not the most efficient tool for achieving environmental targets.

Hoel (1993) examines the need for harmonization of carbon taxes across countries inside an international environmental agreement. He identifies two important problems related to harmonization. Tax harmonization yields equalization of marginal, but not total, cost of emissions across countries and so the rich countries will have smaller total costs. Second, there is the choice of whether the international tax should be added or replace local taxes. Within a formal model Hoel shows that optimally, a uniform international tax should be added to existing taxes. However, the existing free riding incentives will lead countries to the reduction or elimination of local taxes and thus, to a uniform tax which is suboptimal. Thus, it might be appropriate although not optimal to establish a minimum uniform tax for all countries and allow countries to decide whether they add to, or replace existing taxes. Alternatively, a uniform tax supplemented by some reimbursement rules administered by an international body could minimize costs of achieving a given reduction in emissions.

Hoel also discusses the use of tradable emission permits mentioning the advantages of this system over a tax system, namely the simpler institutional arrangements, the ability to be introduced gradually and the direct achievement of an emission reduction. Under ideal conditions, both policy instruments can achieve the efficient allocation but give different allocation of costs between countries. Thus, the reimbursement rules and the initial permit allocation might be based on distributional and not efficiency considerations.

The need to provide incentives for participation in international environmental agreements appears to favor the use of uniform emission standards over more flexible and efficient policy instruments. Some efficiency gains could be recovered if the uniform emission standard policy allows for joint implementation, that is, to permit a number of countries to reach their abatement targets jointly if they agree to do so. The extent to which efficiency can be restored depends on the incentives to participate in joint implementation programs within the agreement. However, it should be noted that the participation to joint implementation programs would be limited, at least at the initial stages, otherwise a market based policy instrument would have been accepted in the first place. Thus, joint implementation can be considered as a supporting element paving the way towards more efficient policies.

5. Conclusions

The present paper reviewed the main developments in the literature that uses a game theoretic framework to analyze international environmental agreements. The main result of this literature is that only small coalitions can be stable, which however could use a variety of policy instruments to expand and sustain broader coalitions. The full coalition could be sustained assuming cooperative behavior and the presence of a supranational authority. However, under the more realistic assumption of non-cooperative strategic behavior and in the absence of a supranational authority, free-riding incentives restrict the size of self-enforcing stable coalitions. Assuming that the members of these small coalitions can credibly commit to certain actions, they can use either their gains from cooperation or other policies (such as links to trade policy, technology transfers and debt policy) to provide participation incentives to other countries.

The analysis of international environmental agreements within a game theoretic framework has been very fruitful. The incorporation of the profitability and stability conditions in non-cooperative strategic behavior and the use of transfer payments and issue linkages are some of the most important contributions. However, there is a number of limitations of this analysis which have mainly to do with the restrictiveness of the assumptions, which we brought forward at the beginning of this review. We believe that future research can contribute substantially in this field either by incorporating more realistic assumptions to the existing framework or by detailing the political process of decision making. On the latter direction, future research could detail the decision making process at the country level (Carraro and Siniscalco (1998) give some ideas in this direction), and/or examine the implementation of the agreement within countries which is assumed as given in the reviewed literature.

References

Barrett, Scott (1992), "International environmental agreements as games." In Rudiger Pethig (editor): Conflicts and Cooperation in Managing Environmental Resources, pages 18-33. Springer, Berlin

Barrett, Scott (1994), "Self-enforcing international environmental agreements." *Oxford Economic Papers*, **46**, 878-894

Barrett, Scott (1997a), "Heterogeneous international environmental agreements." In Carlo Carraro (editor): International Environmental Negotiations: Strategic Policy Issues, chapter 2, pages 9-25. E. Elgar, Cheltenham, UK

Barrett, Scott (1997b), "The strategy of trade sanctions in international environmental agreements." *Resource and Energy Economics*, **19**(4), 345-361

Barrett, Scott (1998), "On the theory and diplomacy of environmental treaty making." *Environmental and Resource Economics*, **11**(3-4), 317-333

Botteon, Michele and Carlo Carraro (1998), "Strategies for environmental negotiation: issue linkage with heterogenous countries." In Nick Hanley and Henk Folmer (editors): Game Theory and the Environment, chapter 9, pages 181-203. E. Elgar, Cheltenham, UK

Breton, Michel Le and Antoine Soubeyran (1998), "The interaction between international environmental and trade policies." In Carlo Carraro (editor): International Environmental Negotiations: Strategic Policy Issues, chapter 8, pages 126-149. E.Elgar, Cheltenham, UK

Carraro Carlo and Domenico Siniscalco (1993), "Strategies for the international protection of the environment." *Journal of Public Economics*, **52,**309-328

Carraro, Carlo and Domenico Siniscalco (1998), "International environmental agreements: Incentives and political economy." *European Economic Review*, **42**, 561-572

Carraro, Carlo and Francesca Moriconi (1997), "International games on climate change control." Paper prepared for the special issue of Natural Resource Modeling on "Environmental Games"
Chander, Parkash Henry Tulkens (1992), "Theoretical foundations of negotiations and cost-sharing in transfrontier pollution problems." *European Economic Review* **36**, 388-398
Chander, Parkash Henry Tulkens (1997), "The core of an economy with multilateral environmental externalities." *International Journal of Game Theory*, **26,** 379-401
De Zeeuw, Art (1998), "International dynamic pollution control." In Nick Hanley and Henk Folmer (editors): Game Theory and the Environment, chapter 12, pages 237-254. E. Elgar, Cheltenham, UK
Ecchia, Giulio , Marco Mariotti (1998a): Coalition formation in international environmental agreements and the role of institutions. *European Economic Review*, **42**, 573-582
Ecchia, Giulio, Marco Mariotti (1998b): "The stability of international environmental coalitions with farsighted countries: some theoretical observations" In Nick Hanley and Henk Folmer (editors): Game Theory and the Environment, chapter 10, pages 30-44. E. Elgar, Cheltenham, UK
Endres, Alfred and Michael Finus (1998), "Renegotiation-proof equilibria in a bargaining game over global emission reductions-does the instrumental framework matter?" In Nick Hanley and Henk Folmer (editors): Game Theory and the Environment, chapter 7, pages 135-164. E. Elgar, Cheltenham, UK
Fankhauser S. and S. Kverndokk (1992), "The global warming game – Simulations of a CO_2 reduction agreement." Working paper GEC WP92-10, CSERGE, university College London
Farrell J. and E. Maskin (1989), "Renegotiation in repeated games." *Games and Economic Behavior*, **1**, 327-60
Finus, Michael and Bianca Rundshagen (1998), "Toward a positive theory of coalition formation and endogenous instrumental choice in global pollution control." *Public Choice*, **96,**145-186
Folmer, Henk, Nick Hanley and Fanny Mibfeldt (1998), "Game-theoretic modeling of environmental and resource problems: an introduction." In Nick Hanley and Henk Folmer (editors): Game Theory and the Environment, chapter 1, pages 1-29. E. Elgar, Cheltenham, UK
Heal, Geoffrey (1994), "Formation of international environmental agreements." In Carlo Carraro (editor): Trade, Innovation, Environment, chapter 3.2, pages 301-322. Kluwer, Dordrecht
Hoel, Michael (1992), "International environmental conventions: the case of uniform reductions of emissions." *Environmental and Resource Economics*, **2,**141-159
Hoel, Michael (1993), "Harmonization of carbon taxes in international climate agreements." *Environmental and Resource Economics*, **3,** 221-231
Hoel, Michael (1994), "Efficient climate policy in the presence of free riders." *Journal of Environmental Economics and Management*, **27**, 259-274
Hoel, Michael and Kerstin Schneider (1997), "Incentives to participate in an international environmental agreement." *Environmental and Resource Economics*, **9**,153-170
Jeppesen T. and P. Andersen (1998), "Commitment and fairness in environmental games.", In Nick Hanley and Henk Folmer (editors): Game Theory and the Environment, chapter 4, pages 65-83. E. Elgar, Cheltenham, UK

Katsoulacos, Yannis (1998), "R&D spillovers, cooperation, subsidies and international agreements." In Carlo Carraro (editor): International Environmental Negotiations: Strategic Policy Issues, chapter 6, pages 97-109. E. Elgar, Cheltenham, UK.

Marriotti, Marco (1997), "A model of agreements in strategic form games." *Journal of Economic Theory*, **74**, p.196-217.

Mohr, Ernst (1995), "International environmental permit trade and debt: The consequences of country sovereignty and cross-default policies." *Review of International Economics*, **3**(1), 1-19

Mohr Ernst and J. Thomas (1998), "Pooling sovereign risks: The case of environmental treaties and international debt." Journal of Development Economics, **55**(1), 153-169

Na, Seong-lin and Hyun Song Shin (1998), "International environmental agreements under uncertainty." *Oxford Economic Papers*, **50**,173-185.

Petrakis, Emmanuel Anastasios Xepapadeas (1996), "Environmental consciousness and moral hazard in international agreements to protect the environment." *Journal of Public Economics*, **60**(1), 95-110

Ploeg, Frederic van der and Aart de Zeeuw (1991), "A differential game of international pollution control." *Systems and Control Letters*, **17,** 409-414

Rabin Mathew (1993), " Incorporating fairness into game theory and economics." *American Economic Review*, **83**(5), 1281-302

Tulkens, Henry (1998), "Cooperation versus free-riding in international environmental and resource problems: two approaches." In Nick Hanley and Henk Folmer (editors): Game Theory and the Environment, chapter 2, pages 30-44. E. Elgar, Cheltenham, UK

Uzawa H (1997), "Global warming as a cooperative game", mimeo

Weiss, Edith Brown (1997), "Strengthening national compliance with international environmental agreements." *Environmental Policy and Law*, **27**(4), 297-303

Xepapadeas, Anastasios and Amalia Yiannaka (1998), "Measuring benefits and damages from carbon dioxide emissions and international agreements to slow down greenhouse warming." In Carlo Carraro (editor): International Environmental Negotiations: Strategic Policy Issues, chapter 9, pages 150-171. E. Elgar, Cheltenham, UK