

Differential Game Approach for International Environmental Agreements with Social Externalities

Lina Mallozzi*, Stefano Patri** , Armando Sacco†

Received 12 November 2015; Accepted 29 March 2016

Abstract In this work we study an N -player differential game, in which positive social externalities affect the payoffs of the players when they make an agreement. We divide the N players in two homogeneous groups, N_1 developed countries and N_2 developing countries. For the latter, we consider a damage-cost function that evolves in time. We imagine the externalities as the possibility that bilateral or multilateral agreements of various nature are by-products of an International Environmental Agreement (IEA). After the determination of emissions solutions, we use the externalities to investigate whether it is possible to have a self-enforcing agreement on pollution emissions in the short run.

Keywords Differential game, self-enforcing agreement, social externality, asymmetric players

JEL classification C72, C73

1. Introduction

A great part of environmental problems, like global warming, depletion of ozone layer or loss of biological diversity, is related to global commons and, for that, requires global policies. During the last three decades, many times countries organized meetings to find an agreement on pollution control. In 1987 in the Canadian city of Montreal, was ratified the Protocol on Substances that Deplete the Ozone Layer. Since 1995 the United Nations organize yearly conferences, Conferences of Parties, within which in 1997 was signed the Kyoto Protocol for the reduction of Green House Gases (GHG), with the objective to contain the global warming.

From an economic point of view, International Environmental Agreements (IEA) lie within the coordination problem class. A natural approach to this kind of problem is game theory. So, there is an extensive literature on this argument, that approaches the problem both as cooperative and non-cooperative games, both as static and dynamic games. The first attempts set the issue in a static context (see e.g. Carraro and Siniscalco 1993; Barrett 1994), and this way is still followed and gives some important contributions. But pollution is an evolving phenomenon, so a dynamic game approach can lead to more explicative results (see e.g., Casino and Rubio 2005; Rubio and Ulph

* University of Naples Federico II, Faculty of Economics, Via Cintia, Monte S. Angelo, Naples 80126, Italy. E-mail: mallozzi@unina.it.

** University of Rome I, Piazzale Aldo Moro 5, Rome 00185, Italy. E-mail: stefano.patri@uniroma1.it.

† Corresponding author. University of Rome I, Piazzale Aldo Moro 5, Rome 00185, Italy. E-mail: armando.sacco@uniroma1.it.

2007; Breton et al. 2010). Both in static and dynamic context, the literature is divided into two streams: cooperative and non-cooperative games. The main focus of the first stream is to contrast the cooperative and non-cooperative solutions and to show the benefits of cooperation. The real question in these games is how to allocate the payoff among players.

The non-cooperative stream starts from the consideration that there is no supranational authority that can force countries to cooperate, thus players choose non-cooperatively whether join or not in a coalition. In this game it is necessary to specify the concept of stability of the coalition. In line with the stable set definition (see von Neumann and Morgenstern 1944), the terms most commonly used are those of internal and external stability, introduced in d'Aspremont et al. (1983). In few words, these two conditions say that a coalition is stable if none of the members has an incentive to defect from cooperation and none of non-members has an incentive to join. So, the two focus points are research of the solutions (emissions or abatement level) and research of the coalition's dimension.

We consider myopic players, that is to say that economic interests are still too strong than environmental concern. It could be a limited point of view. There are relevant examples, like EU, that put into the foreground the control of emissions. But, with the arrival of the new millennium, the economic center of the world has changed his coordinates, and with it the center of the environmental problems. The great challenge now is to include in emissions reduction process those countries that are not considered developed countries, but that give significant contribution to pollution (e.g. the countries called BRICS, Brazil, Russia, India, China, South Africa). We think that it is not realistic to ask those countries to take care of environment for some kind of farsightedness or consider some kind of punishment for those who do not cooperate. So, we want to try to design an IEA that is profitable.

This paper belongs to the non-cooperative games stream (for a recent surveys see Jørgensen et al. 2010; Long 2010). We have a group of developed countries and a group of developing countries. We suppose homogeneity within each group. We also assume that the payoff of cooperators is affected by social externalities that is a function of the coalition's dimension. We recall this idea from Cabon-Dhersin and Ramani (2006). They start from the evidence that, despite theoretic literature on non-cooperative games supports only small coalition, principal agreements are signed by many countries. Thus, their hypothesis is that when countries have to make the decision to join or not an agreement, they consider all possible earnings due to relations with other countries. In this way, they show that in a static game on abatement level, with N symmetric countries, it is possible to have a grand coalition. Our purpose is to verify whether their results are true also in a dynamic contest with asymmetric players. We work on emissions instead on abatement levels (for a discussion on the duality between emission and abatement level see Diamantoudi and Sartzetakis 2006; Finus 2001) and include the social externalities in an N -player differential game in the framework of Masoudi and Zaccour (2013). Thus, we have two different treatments for environmental concern of the two kinds of players. The idea is that developing countries have an environmental damage-cost that is not full from the outset, but increases in time with

the increasing of their cumulative revenue.

Our main focus is on two points:

1. To find the feedback-Nash equilibrium emissions both for cooperators and defectors, and
2. to evaluate the size and the composition of the coalition by using the concept of internal and external stability.

The rest of the paper is organized as follows: In Section 2 we introduce the model. In Section 3 we characterize the cooperative and non-cooperative emission solutions. In Section 4 we discuss the stability of the coalition and show some numerical results and Section 5 concludes.

2. The Model

2.1 Social externalities

Why a country should join an agreement for pollution control? This is the main issue at which environmental economics literature tried to respond in the last decades. Actually, two real issues should be considered. The first is the need to involve developing countries in emissions' reduction process. The second is to consider the IEAs within the complexity of a large number of relationships between countries. The mechanisms most used in economic literature to reach a large coalition are *transfer scheme strategies* (see e.g., Fuentes-Albero and Rubio 2010; Pavlova and de Zeeuw 2013) and *trigger strategies* (see e.g., Hoel and Schneider 1997). The idea of a transfer scheme is that signatories use the gain from cooperation to convince defectors to join in the coalition. On the other side, in the case of trigger strategies, the assumption is that signatories have the power to punish the defectors.

A further way is the *issue linkage* literature (see e.g., Botteon and Carraro 1994; Le Breton and Soubeyran 1997; Hübler and Finus 2013), in which the IEA is linked with another agreements, that could be a R&D, or trade or another economic issues. Considering the globalization of relations between countries, we think that transfer scheme and trigger strategies are not credible mechanisms to enlarge an environmental coalition. Issue linkage could be a suitable way to model an IEA, but we think that specifying a kind of side agreement does not allow to take into account several other connection between countries.

Our approach to solve the problem, is to assume that social externalities affect the pay-off of players that decide to join in a coalition. From a mathematical point, we assume that a strictly positive function is added to the pay-off of signatories, and this function does not depend on emission, but only on the number of players in coalition. The principal reason of that choice lies in the great flexibility of the possible interpretation of what this externalities represent. We only request that when countries decide whether to join a coalition, they consider the possible benefits deriving from being in relationships with other countries. A classical example is Russia, that ratified the Kyoto protocol with the hope to have more consideration when its entry in World Trade Organization (WTO) would have been voted. We can also imagine that collateral to environmental agreement, other agreements, like trading, R&D ecc., could be signed

(like in issue linkage literature). The immateriality of this social externalities brings with it a certain degree of vagueness, because within this concept we include all the possible networks that countries could establish. Nevertheless, we think that the loss of descriptive power is acceptable, compared to the great flexibility that we gain.

2.2 Functional forms

We consider an N -player differential game, assuming the world divided in two types of countries. So, we have a total number of players $N = N_1 + N_2$, in which N_1 are developed countries and N_2 are developing countries (N , N_1 and N_2 are integer numbers). We use 1 to denote developed countries and 2 to denote developing countries. As usual, we proceed backward.

We first assume that a number k of players join the agreement, while the rest stays out. In particular we take $k = n_1 + n_2$, where n_1 are developed countries and n_2 are developing countries, and k, n_1, n_2 are integer numbers. The assumption that some players that act cooperatively, while other players act non-cooperatively, belongs into partial cooperative games (for a discussion see Mallozzi and Tijs 2008, 2009; Chakrabarti et al. 2011). According to this assumption, we first solve the emission game and then we use the optimal emissions to find the numbers n_1 and n_2 that satisfy the stability conditions.

First of all, due to the fact that emissions are by-product of industrial activities, and assuming that the function which relates emissions and production are smooth and invertible, we can express the production for country i as a function of emission levels. Denoting it with $f_i(e_i)$, being e_i the emission of country i , and assuming that it is a concave and increasing function, that is a standard assumption in literature (see, e.g., Finus 2001; Rubio and Casino 2005; Diamantoudi and Sartzetakis 2006), we can define the productions as

$$f_i(e_i) = \alpha_i e_i - \frac{1}{2} e_i^2,$$

where $\alpha_i > 0$, so that $f_i(e_i)$ is positive for all suitable value of emissions.

Moreover, our hypothesis is that developed countries have a higher degree of interest in environmental issues with respect to developing countries, both for economic and historical motivations. In the end, this is the same approach of the Kyoto protocol.

The point is that a developing country needs to improve its infrastructure, increasing per capita wealth, life expectancy, instruction level, etc. In this context, environment is a “luxury good”. In addition, the actual level of stock pollutant cannot be attributed to developing countries.

So, we have a different degree of internalization of the environmental damage-cost, that is given by a different definition of the damage-cost functions. Being $S(t)$ the stock of pollutant at time t , we denote by $D_1(S)$ the cost for developed countries and we assume that it is full from the outset. We choose a linear function of stock of pollutant, that is a not uncommon choice (see Hoel and Schneider 1997; Breton et al. 2010) and it is supported by some empirical estimation (see Labriet and Loulou 2003). In the end, the difference between linear and a more realistic quadratic damage-cost

function should be only quantitative, but not qualitative. Then, we assume

$$D_1(S) = \beta_1 S,$$

where $\beta_1 > 0$, so that the damage-cost function is an increasing function of stock of pollutant.

On the other side, for developing countries the full damage-cost is related to the achievement of a preset threshold in terms of cumulative discounted revenues, denoted by $Y_2(t)$. The idea of linking income and environmental quality is not new in literature, there is the well-known environmental Kutznets curve (EKC) and some works, like Shafik and Bandyopadhyay (1992), that support the empirical consistence of this hypothesis. Then, denoting by ρ the rate of time preference, that we assume common to all players, we have $Y_2(t) = \int_0^t f_i(e_i(z))e^{-\rho z}dz$. Given that relation, we can define the time T as the instant ($T > 0$) at which a country of type 2 start to fully account the damage-cost, that is the time at which

$$\int_0^T f_i(e_i(t))e^{-\rho t}dt = \overline{Y}_2$$

is verified, where \overline{Y}_2 is the threshold chosen.

Then, for players 2 the damage-cost function is described for any t in two intervals:

$$\begin{cases} d_2(S, t) = \frac{t}{T} \gamma \beta_2 S, & \forall t < T, \\ D_2(S) = \beta_2 S, & \forall t \geq T, \end{cases}$$

where $\beta_2 > 0$. Moreover, we suppose $\gamma \in \{0, 1\}$. The case $\gamma = 0$ is equivalent to say that players of type 2 do not allow for pollution at all, until they reach the threshold \overline{Y}_2 (that's the spirit of Kyoto protocol). If $\gamma = 1$, we are in the case of gradual internalization of damage cost. The stock of pollutant $S(t)$ is solution of the following differential equation:

$$\dot{S}(t) = \mu \left(\sum_{i=1}^N e_i(t) \right) - \delta S(t), \quad S(0) = S_0,$$

where μ is a positive scaling parameter and δ is the natural rate of absorption of pollution. Here S_0 is the initial value of the pollution.

From now on, we will skip the time argument if there is no risk of ambiguity. We can introduce the payoff functions, that we denote by w_i , $i = 1, 2$, given by

$$w_1 = \int_0^\infty (f_1(e_1) - D_1(S))e^{-\rho t}dt,$$

$$w_2 = \int_0^T (f_2(e_2) - d_2(S, t))e^{-\rho t}dt + \int_T^\infty (f_2(e_2) - D_2(S))e^{-\rho(t-T)}dt.$$

The last function that we need to characterize is the social externalities, that we denote by $Ext(n_1, n_2)$. As in Cabon-Dhersin and Ramani (2006), we make the assumption that

it is a positive linear function of the coalition size, as follows

$$Ext(n_1, n_2) = s_1 n_1 + s_2 n_2,$$

where we assume the marginal externality s_i to be positive for $i = 1, 2$. Obviously, with n_i we state the number of players of kind i that join in coalition, precisely $n_i \in \{0, 1, \dots, N_i\}$, $i = 1, 2$.

3. Emission Solution

In this section we characterize the emission solutions both for signatories and defectors. We suppose that we have a set C of signatory players, with n_1 developed countries and n_2 developing countries and a set of defectors, that we denote with NC . Then, the cardinality of the two sets is given by $|C| = n_1 + n_2$ and $|NC| = N_1 - n_1 + N_2 - n_2$. As usual, every player $j \in NC$ maximizes his own welfare, while players in C maximize the joint welfare. Due to homogeneity within groups, we just have to find four emission solutions: two for signatories (called e_1^C and e_2^C), and two for the defectors (called e_1^{NC} and e_2^{NC}).

Thus, the problem for defectors is:

$$\max_{e_1} \int_0^\infty (f_1(e_1) - D_1(S)) e^{-\rho t} dt, \quad (1)$$

and

$$\max_{e_2} \int_0^{T^{NC}} (f_2(e_2) - d_2(S, t)) e^{-\rho t} dt + \int_{T^{NC}}^\infty (f_2(e_2) - D_2(S)) e^{-\rho(t-T^{NC})} dt \quad (2)$$

$$\text{s.t. } \dot{S}(t) = \mu \left(\sum_1^N e_i(t) \right) - \delta S(t), \quad S(0) = S_0.$$

For signatories, the joint maximization is:

$$\max_{e_1, e_2} \int_0^{T^C} (n_1 f_1(e_1) + n_2 f_2(e_2) - n_1 D_1(S) - n_2 d_2(S, t) + Ext(n_1, n_2)) e^{-\rho t} dt + \int_{T^C}^\infty (n_1 f_1(e_1) + n_2 f_2(e_2) - n_1 D_1(S) - n_2 D_2(S) + Ext(n_1, n_2)) e^{-\rho(t-T^C)} dt, \quad (3)$$

$$\text{s.t. } \dot{S}(t) = \mu \left(\sum_1^N e_i(t) \right) - \delta S(t), \quad S(0) = S_0.$$

In the maximization problems we call T^{NC} and T^C the instants of time at which a developing country achieves the threshold to become developed, respectively in the cases of defector and signatory. It seems clear from the optimization problems that we consider a feedback emission game, but an open-loop membership game.

3.1 Emissions of defectors

To solve the problem, we use the dynamic programming method. We proceed backward, solving first the problem on $[T^{NC}, \infty)$. So, we have to solve first:

$$\begin{aligned} \max_{e_i} \int_{T^{NC}}^{\infty} \left(\alpha_i e_i - \frac{1}{2} e_i^2 - \beta_i S \right) e^{-\rho(t-T^{NC})} dt, \\ \text{s.t. } \dot{S}(t) = \mu \left(n_1 e_1^C(t) + n_2 e_2^C(t) + (N_1 - n_1) e_1^{NC}(t) + (N_2 - n_2) e_2^{NC}(t) \right) - \delta S(t), \end{aligned}$$

where $S(0) = S_0$. Using Hamilton-Jacobi-Bellman (HJB) equations, we find the solution:

$$e_i^{NC}(t) = \alpha_i - \mu \frac{\beta_i}{\rho + \delta}, \quad i = 1, 2. \quad (4)$$

Since the functional forms for players of type 1 is the same in the entire period, we have that the emissions of developed countries are in $[0, T^N]$:

$$e_1^{NC}(t) = \alpha_1 - \mu \frac{\beta_1}{\rho + \delta}.$$

For developing countries we have to consider the different damage-cost function, so the problem is:

$$\begin{aligned} \max_{e_2} \int_0^{T^{NC}} \left(\alpha_2 e_2 - \frac{1}{2} e_2^2 - \gamma \frac{t}{T^{NC}} \beta_2 S \right) e^{-\rho t} dt, \\ \text{s.t. } \dot{S}(t) = \mu \left(n_1 e_1^C(t) + n_2 e_2^C(t) + (N_1 - n_1) e_1^{NC}(t) + (N_2 - n_2) e_2^{NC}(t) \right) - \delta S(t), \end{aligned}$$

where $S(0) = S_0$. In this case HJB equation leads us to the emissions, for $t \in [0, T^N]$:

$$\begin{aligned} e_2^{NC}(t) = \alpha_2 - \mu \frac{\beta_2}{T^{NC}(\rho + \delta)^2} \left[\gamma(1 + t(\rho + \delta) - e^{(\rho + \delta)(t - T^{NC})}) + \right. \\ \left. + T^{NC}(1 - \gamma)(\rho + \delta)e^{(\rho + \delta)(t - T^{NC})} \right]. \end{aligned} \quad (5)$$

3.2 Emissions of signatories

As the previous case, we proceed backwards to find the optimal emissions of signatories. First of all we solve the problem in $[T^C, \infty)$:

$$\begin{aligned} \max_{e_1, e_2} \int_{T^C}^{\infty} \left[n_1 \left(\alpha_1 e_1 - \frac{1}{2} e_1^2 \right) + n_2 \left(\alpha_2 e_2 - \frac{1}{2} e_2^2 \right) - n_1 \beta_1 S - n_2 \beta_2 S + s_1 n_1 + s_2 n_2 \right] e^{-\rho(t-T^C)} dt, \\ \text{s.t. } \dot{S}(t) = \mu \left(n_1 e_1^C(t) + n_2 e_2^C(t) + (N_1 - n_1) e_1^{NC}(t) + (N_2 - n_2) e_2^{NC}(t) \right) - \delta S(t), \end{aligned}$$

where $S(0) = S_0$.

We want to highlight that the externalities has no effects on the determination of the emissions. Using HJB, we determine the solution for signatory countries as:

$$e_i^C(t) = \alpha_i - \mu \frac{n_1 \beta_1 + n_2 \beta_2}{\rho + \delta}, \quad i = 1, 2. \quad (6)$$

In the period $[0, T^C]$, the different damage-cost function implies that the optimal solutions solve:

$$\begin{aligned} \max_{e_1, e_2} \int_0^{T^C} & \left[n_1 \left(\alpha_1 e_1 - \frac{1}{2} e_1^2 \right) + n_2 \left(\alpha_2 e_2 - \frac{1}{2} e_2^2 \right) - n_1 \beta_1 S - n_2 \gamma \frac{t}{T^C} \beta_2 S + s_1 n_1 + s_2 n_2 \right] e^{-\rho t} dt, \\ \text{s.t. } \dot{S}(t) = & \mu \left(n_1 e_1^C(t) + n_2 e_2^C(t) + (N_1 - n_1) e_1^{NC}(t) + (N_2 - n_2) e_2^{NC}(t) \right) - \delta S(t), \end{aligned}$$

where $S(0) = S_0$. The optimal emissions are given by:

$$\begin{aligned} e_i^C(t) = \alpha_i - \mu \frac{n_1 \beta_1}{\rho + \delta} - \mu \frac{n_2 \beta_2}{T^C(\rho + \delta)^2} & \left[\gamma(1 + t(\rho + \delta) - e^{(\rho + \delta)(t - T^C)}) + \right. \\ & \left. + T^C(1 - \gamma)(\rho + \delta)e^{(\rho + \delta)(t - T^C)} \right], \end{aligned} \quad (7)$$

where $i = 1, 2$. Nevertheless, we will see that externalities will effect the payoffs of the players.

4. Stability

In order to approach the stability analysis, we use the concept of self-enforcing agreements. The idea is due to d'Aspremont et al. (1983). They use this concept to study the stability of a cartel, but in several places, it is also used to discuss the stability of environmental agreements. We want to highlight that these conditions are more stringent and there are different papers that try to propose different ways to face the problem (see Finus 2003; Eyckmans and Finus 2006). The basic idea is that a coalition is stable if no one inside has an incentive to defect and no one outside has an incentive to join in. So, called w_i the pay-off of a player i , a coalition of k players is stable if

$$w_i^C(k) \geq w_i^{NC}(k-1), \quad w_j^{NC}(k) \geq w_j^C(k+1)$$

are verified $\forall i \in C$ and $\forall j \in NC$. First condition is called *internal* stability, while the second is called *external* stability.

In our case, having two different types of players, we need to adapt the definition, requiring that internal and external stability are verified both for developed and developing countries. So, we have to find the values n_1 and n_2 that solve this system of four

inequalities:

$$\begin{cases} w_1^C(n_1, n_2) - w_1^{NC}(n_1 - 1, n_2) \geq 0, \\ w_2^C(n_1, n_2) - w_2^{NC}(n_1, n_2 - 1) \geq 0, \\ w_1^{NC}(n_1, n_2) - w_1^C(n_1 + 1, n_2) \geq 0, \\ w_2^{NC}(n_1, n_2) - w_2^C(n_1, n_2 + 1) \geq 0. \end{cases} \quad (8)$$

As the agreements are revised periodically, we focus our analysis only on the short term, that is the period $[0, T]$. Unfortunately, we are not able to solve the system analytically, due to the complexity and to the nonlinearity of the functions. So, we analyze the problem from a numerical point of view. We make some simulations using the software Wolfram Mathematica. The first step is the calibration of parameters in a benchmark model, that we summarize in Table 1.

Table 1. Benchmark parameters for developed (Type 1) and developing (Type 2) countries

	i=1	i= 2	
α_i	3.38	2.32	$\mu = 0.64$
β_i	0.0031	0.0048	$\delta = 0.0083$
s_i	0.08	0.1	$\rho = 0.01$

To calibrate the parameters of production and damage-cost functions we use the World Bank data set of emissions, expressed in kg per purchasing power parity (PPP) of Gross Domestic Product (GDP). We use the aggregate data of the developed countries and the upper middle income countries, within that we have all the developing countries with a significant industrial structure (e.g., BRICS). Based on this division of the World Bank, we set $N = 10$, with $N_1 = 6$ and $N_2 = 4$.

Moreover our choice of δ and μ is based on Nordhaus (1993). The parameters s_1 and s_2 in the function $Ext(n_1, n_2)$ are chosen sufficiently small not to overestimate the effect of the social externalities. Because by our simulation the starting level of pollution seems to be not relevant for the stability, we can assume $S_0 = 0$. We make two different studies, one with $\gamma = 0$ and one with $\gamma = 1$. In the following, we present only the case latter case, because it is more interesting for the sensitivity analysis and the case $\gamma = 0$ does not present substantial differences.

The first values that we have to compute are T^{NC} and T^C , taking every time a suitable value for \bar{Y}_2 , which depends on the production function. As we expected, we have always $T^{NC} \leq T^C$, so we focus our simulation on the interval $[0, T^{NC}]$. After that, we solve the differential equations for $S(t)$ in the different configurations required by stability conditions. Having all the elements we need, we can proceed with the simulation of stability. We evaluate the 32 possible coalitions, due to the different combinations of variables n_1 and n_2 .

The benchmark model gives an univocal result: the only stable coalition is the grand coalition. What we observed is that all possible coalitions are internally stable, but no one is externally stable, that means that all players have an incentive to join

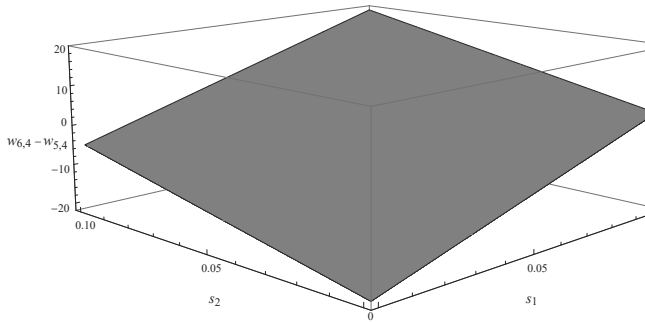


Figure 1. Internal stability for developed countries as function of s_1 and s_2 , with $n_1 = 6$ and $n_2 = 4$

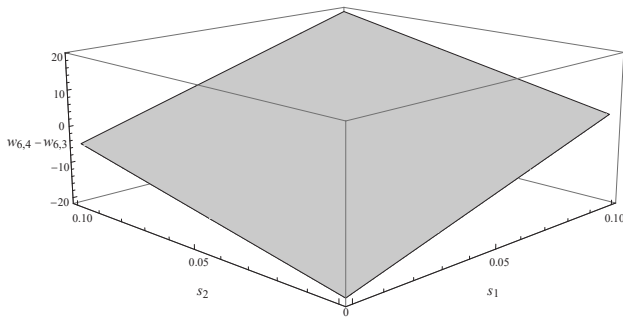


Figure 2. Internal stability for developing countries as function of s_1 and s_2 , with $n_1 = 6$ and $n_2 = 4$

the agreement. Another thing we learn from this model is that the grand coalition stops being stable only if the parameters s_1 and s_2 converge to zero. In Figures 1 and 2 we represent the internal stability conditions of the system (8), as function of the two parameters s_1 and s_2 .¹ We can see that the coalition is unstable if there are no social externalities or if at least one of the two groups gives little weight to non environmental possibilities given by the agreement. We make the same simulations on smaller coalition and the surfaces that we obtain have a shape very close to that of Figures 1 and 2.

Going forward with our sensitivity analysis, we pass to the parameters of the production function, α_1 and α_2 . The results of the simulations say that, while there is no effect due to a change in α_1 , the parameter α_2 influences the stability only in the measure of change of T^{NC} , but the grand coalition is still the only stable.

We obtain different results if we test the sensitivity of the model to variations in damage-cost parameters, β_1 and β_2 . In this case an increment of the vulnerability to

¹ Clearly, we have not external stability conditions in case of Grand Coalition. We make some simulation on external stability for smaller coalition, and we found that the qualitative results are symmetric to these.

Table 2. Stable coalitions for simultaneous variation of β_1 and β_2

	Marginal costs sensitivity		
	$\beta_1 = 0.0062$	$\beta_1 = 0.0016$	$\beta_1 = 0.0062$
	$\beta_2 = 0.0024$	$\beta_2 = 0.0096$	$\beta_2 = 0.0096$
Stable coalitions	(5, 0)	(0, 4)	(0, 4)
			(5, 0)
			(1, 3)
			(2, 2)
			(4, 1)

the damage of pollution can make large coalitions unstable. The first test we make in this sense in taking a double value of β_1 , that is $\beta_1 = 0.0062$, given the values of the other parameters. This increment brings to a situation of homogeneity: we still have a unique stable solution, but now the coalition is formed only by developed countries, in fact the solution couple is $(n_1, n_2) = (5, 0)$.

The next step is the evaluation of an increment in the value of β_2 , so we take $\beta_2 = 0.0096$, given the other parameters as in the benchmark model. The results are symmetric to the previous one: We have only one stable and homogeneous coalition, but in this case it is the one formed only by developing countries, that is $(n_1, n_2) = (0, 4)$. The difference is that this coalition consists of all the developing countries, while above we have a coalition formed by all but one developed countries.

We also test the effects of a simultaneous variations of parameters β_1 and β_2 . We first try to balance out the growth of one parameters with a reduction of the other. In the first column of the Table 2, we show the results on the stability in the case that β_1 is doubled and β_2 is the half. As we can see, reduction of the latter doesn't bring some compensation in terms of stable coalitions and the only stable coalition is still $(n_1, n_2) = (5, 0)$. The same thing happens in the second column of Table 2, in which β_1 decreases and β_2 increases. Also in this case, the only stable coalition is $(n_1, n_2) = (0, 4)$. Then, decreasing one of the marginal damage costs can not compensate in any an increase of the other damage cost. More interesting the last column, in which both parameters are doubled. In this case we have an enlargement of the set of stable solutions, which now includes also the mixed coalitions (1, 3), (2, 2) and (4, 1).

5. Conclusions

To develop efficient policies on pollution control, stability of an International Environmental Agreement is the key. In this work we investigated a non-cooperative N -player differential game, in which we divided the world in two types of players, developed and developing countries. For the latter we assumed an evolving damage-cost function, taking into account the particular issues of this countries. Our main contribution is to recall the idea of social externalities, investigated by Cabon-Dhersin and Ramani (2006) in a static context with symmetric players, and to verify its validity in a more

general and realistic context.

First of all we characterized the emission solutions both for signatories and defectors, assuming as a date the size of the coalition. After that, we used these solutions to study the stability of the agreement. Due to the non-linearity and the complexity of the model, we analyzed the problem with numerical simulations.

Our benchmark model shows that we can have just one stable coalition, namely a coalition of all players. Moreover, this coalition become unstable only if the parameters s_1 and s_2 converge to zero. Nevertheless, there is a significant sensitivity to the parameters β_1 and β_2 , so if they increase, we have smaller stable coalitions. To conclude, we want to highlight two possible ways to extend the model. The first point is to consider a feedback game also for the membership, including the possibility that players join the agreement in different times. The second one is to consider a higher degree of asymmetry, going beyond the assumption of homogeneity within the two groups of countries.

Acknowledgment We would like to thank Fabio Di Dio (SOGEI SPA) for his valuable comments and suggestions.

References

- Barrett, S. (1994). Self-Enforcing International Environmental Agreements. *Oxford Economic Papers*, 46, 878–894.
- Botteon, M. and Carraro, C. (1994). Strategies for Environmental Negotiations: Issue Linkage with Heterogeneous Countries. In Hanley, N., Folmer, H. (eds.), *Game Theory and Environment*. Cheltenham, Edward Elgar, 181–203.
- Breton, M., Sbragia, L. and Zaccour, G. (2010). A Dynamic Model for International Environmental Agreements. *Environmental and Resource Economics*, 45, 25–48.
- Cabon-Dhersin, M. L. and Ramani, S. V. (2006). Can Social Externalities Solve the Small Coalition Puzzle in International Environmental Agreements?. *Economics Bulletin*, 17(4), 1–8.
- Carraro, C. and Siniscalco, S. (1993). Strategies for International Protection of the Environment. *Journal of Public Economics*, 52, 309–328.
- Casino, B. and Rubio, S. J. (2005). Self-Enforcing International Environmental Agreements with a Stock Pollutant. *Spanish Economic Review*, 7, 89–109.
- Chakrabarti, S., Gilles, R. P. and Lazarova, E. A. (2011). Strategic Behaviour under Partial Cooperation. *Theory and Decision*, 71(2), 175–193.
- d'Aspremont, C., Jacquemin, A., Gabszewicz, J. J. and Weymark, J. A. (1983). On the Stability of Collusive Price Leadership. *The Canadian Journal of Economics*, 16, 17–25.
- Diamantoudi, E. and Sartzetakis, E. S. (2006). Stable International Environmental Agreements: An Analytical Approach. *Journal of Public Economic Theory*, 8, 247–263.

- Eyckmans, J. and Finus, M. (2006). A Coalition Formation in a Global Warming Game: How the Design of Protocols Affects the Success of Environmental Treaty-Making. *Natural Resource Modeling*, 19(3), 323–358.
- Finus, M. (2001). *Game Theory and International Environmental Cooperation*. Cheltenham, Edward Elgar.
- Finus, M. (2003). Stability and Design of International Environmental Agreements: The Case of Transboundary Pollution. In Folmer, H. and Tietenberg, T. (eds.), *The International Yearbook of Environmental and Resource Economics 2003/2004: A Survey of Current Issues*. Cheltenham, Edward Elgar, 82–158.
- Fuentes-Albero, C. and Rubio, S. (2010). Can International Environmental Cooperation Be Bought? *European Journal of Operational Research*, 202, 255–264.
- Hoel, M. and Schneider, K. (1997). Incentives to Participate in an International Environmental Agreement. *Environmental and Resource Economics*, 9, 153–170.
- Hübler, M. and Finus, M. (2013). Is the Risk of North-South Technology Transfer Failure an Obstacle to a Cooperative Climate Change Agreement? *International Environmental Agreements: Politics, Law and Economics*, 13(4), 461–479.
- Jørgensen, S., Martín-Herrán, G. and Zaccour, G. (2010). Dynamic Games in the Economics and Management of Pollution. *Environmental Modeling and Assessment*, 15(6), 433–467.
- Labriet, M. and Loulou, R. (2003). Coupling Climate Changes and GHG Abatement Costs in a Linear Programming Framework. *Environmental Modeling and Assessment*, 8, 261–274.
- Le Breton, M. and Souberyan, A. (1997). The Interaction between International Environmental and Trade Policies. In C. Carraro (ed.), *International Environmental Negotiations – Strategic Policy Issues*. Cheltenham, Edward Elgar, 126–149.
- Long, N. V. (2010). *Dynamic Games in Economics: A Survey*. Singapore, World Scientific Publishing Company.
- Mallozzi, L. and Tijs, S. (2008). Conflict and Cooperation in Symmetric Potential Games. *International Game Theory Review*, 10(3), 1–12.
- Mallozzi, L. and Tijs, S. (2009). Coordinating Choice in Partial Cooperative Equilibrium. *Economics Bulletin*, 29(2), 1467–1473.
- Masoudi, N. and Zaccour, G. (2013). A Differential Game of International Pollution Control with Evolving Environmental Costs. *Environment and Development Economics*, 18(6), 680–700.
- Nordhaus, W. D. (1993). Rolling the Dice: An Optimal Transition Path for Controlling Greenhouse Gases. *Resource and Energy Economics*, 15(1), 27–50.
- Pavlova, Y. and de Zeeuw, A. (2013). Asymmetries in International Environmental Agreements. *Environment and Development Economics*, 18, 51–68.

Rubio, S. J. and Ulph, A. (2007). An Infinite-Horizon Model of Dynamic Membership of International Environmental Agreements. *Journal of Environmental Economics and Management*, 54, 296–310.

Shafik, N. and Bandyopadhyay, S. (1992). Economic Growth and Environmental Quality: Time Series and Cross-Country Evidence. Policy, research working papers No. 904. World development report. Washington, DC, World Bank.

von Neumann, J. and Morgenstern, O. (1944). *Theory of Games and Economic Behaviour*. Princeton, Princeton University Press.

Appendix

A. Emissions of defectors

To solve the problem, we use dynamic programming. For given time T^{NC} , we proceed backward, solving first the problem on $[T^{NC}, \infty)$. We have to solve first:

$$\begin{aligned} \max_{e_i^{NC}} \int_{T^{NC}}^{\infty} \left(\alpha_i e_i^{NC} - \frac{1}{2} (e_i^{NC})^2 - \beta_i S \right) e^{-\rho(t-T^{NC})} dt, \\ \text{s.t. } \dot{S}(t) = \mu \left(\sum_1^N e_i(t) \right) - \delta S(t), \quad S(0) = S_0. \end{aligned}$$

If we denote with $v(t, S)$ the value function of the problem, we can write the Hamilton-Jacobi-Bellman (HJB) equation as

$$-v_t = \max_{e_i^{NC}} \left\{ \left(\alpha_i e_i - \frac{1}{2} (e_i^{NC})^2 - \beta_i S \right) e^{-\rho t} + v_S \left(\mu \sum_{j=1}^N e_j - \delta S \right) \right\}. \quad (9)$$

Solving the first order conditions in (9), we obtain an expression for the optimal emissions

$$e_i^{NC}(t) = \alpha_i + \mu v_S e^{\rho t}.$$

Let us assume that value function $v(t, S)$ is linear in S :

$$v(t, S) = (KS + L)e^{-\rho t},$$

with partial derivatives $v_t = -\rho(KS + L)e^{-\rho t}$ and $v_S = Ke^{-\rho t}$. So, emissions for a player i outside coalition are given by

$$e_i^{NC}(t) = \alpha_i + K\mu.$$

In order to find an expression for the parameter K , we substitute v_t , v_S and e_i inside (9):

$$\rho(KS + L)e^{-\rho t} = \left[\alpha_i(\alpha_i + K\mu) - \frac{1}{2}(\alpha_i + K\mu)^2 - \beta_i S \right] e^{-\rho t} + K \left(\mu \sum_{j=1}^N e_j - \delta S \right) e^{-\rho t}$$

$$\rho(KS + L) = \alpha_i^2 + \alpha_i K\mu - \frac{1}{2}\alpha_i^2 - \frac{1}{2}K^2\mu^2 - \alpha_i K\mu - \beta_i S + K\mu \sum_{j=1}^N e_j - K\delta S$$

$$\rho KS + \rho L = -(\beta_i + K\delta)S + \frac{1}{2}\alpha_i^2 - \frac{1}{2}K^2\mu^2 + K\mu \sum_{j=1}^N e_j.$$

By the principle of identity of polynomials, we can write the equation

$$\rho K = -\beta_i - K\delta,$$

from which we have that $K = -\frac{\beta_i}{\rho+\delta}$.

Finally, we have the expression of the emissions of defectors

$$e_i^{NC}(t) = \alpha_i - \mu \frac{\beta_i}{\rho + \delta}, \quad i = 1, 2. \quad (10)$$

We proceed now to find the optimal emissions for $t \in [0, T^{NC}]$. Since the functional forms for players of type 1 is the same in the entire period, we have that the emissions of developed countries are the same in $[0, T^{NC}]$:

$$e_1^{NC}(t) = \alpha_1 - \mu \frac{\beta_1}{\rho + \delta}, \quad \forall t \geq 0.$$

For developing countries we have to consider the different damage-cost function, so the problem is:

$$\begin{aligned} \max_{e_2^{NC}} \int_0^{T^{NC}} & \left(\alpha_2 e_2^{NC} - \frac{1}{2} (e_2^{NC})^2 - \gamma \frac{t}{T^{NC}} \beta_2 S \right) e^{-\rho t} dt, \\ \text{s.t. } \dot{S}(t) &= \mu \left(\sum_1^N e_i(t) \right) - \delta S(t), \quad S(0) = S_0. \end{aligned}$$

In this case the HJB equation is given by

$$-v_t = \max_{e_2} \left\{ \left(\alpha_2 e_2 - \frac{1}{2} e_2^2 - \gamma \frac{t}{T^{NC}} \beta_2 S \right) e^{-\rho t} + v_S \left(\mu \sum_{j=1}^N e_j - \delta S \right) \right\}. \quad (11)$$

As usual, we derive the first order conditions from maximization in (11)

$$(\alpha_2 - e_2) e^{-\rho t} + v_S \mu = 0 \Rightarrow e_2^{NC}(t) = \alpha_2 + v_S \mu e^{\rho t}.$$

For the value function $v(t, S)$ we assume

$$v(t, S) = [x(t)S + y(t)]e^{-\rho t},$$

whose partial derivatives are:

$$v_t = [(x'(t) - \rho x(t))S + y'(t) - \rho y(t)]e^{-\rho t} \quad \text{and} \quad v_S = x(t)e^{-\rho t}.$$

The optimal emissions are given by

$$e_2^{NC}(t) = \alpha_2 + \mu x(t).$$

Now, we have to substitute v_t , v_S and e_2 inside (11)

$$-[(x'(t) - \rho x(t))S + y'(t) - \rho y(t)]e^{-\rho t} = \left[\alpha_2(\alpha_2 + \mu x(t)) - \frac{1}{2}(\alpha_2 + \mu x(t))^2 - \gamma \frac{t}{T^{NC}} \beta_2 S \right] e^{-\rho t} + x(t) \left(\mu \sum_{j=1}^N e_j - \delta S \right) e^{-\rho t}.$$

Rearranging with respect to the principle of identity of polynomials, we can write the differential equation

$$\begin{cases} -x'(t) + (\rho + \delta)x(t) = -\gamma \frac{t}{T^{NC}} \beta_2, \\ x(T^{NC}) = -\frac{\beta_2}{\rho + \delta}. \end{cases}$$

The solution $x(t)$ is given by

$$x(t) = -\frac{\beta_2}{T^{NC}(\rho + \delta)^2} \left[\gamma(1 + t(\rho + \delta) - e^{(\rho + \delta)(t - T^{NC})}) + e^{(\rho + \delta)(t - T^{NC})} T^{NC}(\rho + \delta)(1 - \gamma) \right],$$

and leads us to the developing countries' emissions, for $t \in [0, T^{NC}]$:

$$e_2^{NC}(t) = \alpha_2 - \mu \frac{\beta_2}{T^{NC}(\rho + \delta)^2} \left[\gamma(1 + t(\rho + \delta) - e^{(\rho + \delta)(t - T^{NC})}) + T^{NC}(1 - \gamma)(\rho + \delta)e^{(\rho + \delta)(t - T^{NC})} \right]. \quad (12)$$

B. Emissions of signatories

As for the defectors' case, we assume as known the time T^C and we proceed backward. First of all solve the problem for $t \in [T^C, \infty)$:

$$\max_{e_1^C, e_2^C} \int_{T^C}^{\infty} \left[n_1 \left(\alpha_1 e_1^C - \frac{1}{2}(e_1^C)^2 \right) + n_2 \left(\alpha_2 e_2^C - \frac{1}{2}(e_2^C)^2 \right) - n_1 \beta_1 S - n_2 \beta_2 S + s_1 n_1 + s_2 n_2 \right] e^{-\rho(t - T^C)} dt,$$

$$\begin{aligned} \text{s.t. } \dot{S}(t) &= \mu \left(n_1 e_1^C(t) + n_2 e_2^C(t) + (N_1 - n_1) e_1^{NC}(t) + (N_2 - n_2) e_2^{NC}(t) \right) - \delta S(t), \\ S(0) &= S_0. \end{aligned}$$

We want to highlight that the externalities have no effects on the determination of the emissions. Let $v(t, S)$ be the value function, the HJB equation is

$$-v_t = \max_{e_1^C, e_2^C} \left\{ \left[n_1 \left(\alpha_1 e_1^C - \frac{1}{2} (e_1^C)^2 \right) + n_2 \left(\alpha_2 e_2^C - \frac{1}{2} (e_2^C)^2 \right) - n_1 \beta_1 S - n_2 \beta_2 S \right] e^{-\rho t} + v_S \left[\mu \left(n_1 e_1^C + n_2 e_2^C + (N_1 - n_1) e_1^{NC} + (N_2 - n_2) e_2^{NC} \right) - \delta S \right] + (s_1 n_1 + s_2 n_2) e^{-\rho t} \right\}. \quad (13)$$

The first order conditions in (13) are given by

$$\begin{aligned} n_1 [(\alpha_1 - e_1^C) e^{-\rho t} + \mu v_S] &= 0, \\ n_2 [(\alpha_2 - e_2^C) e^{-\rho t} + \mu v_S] &= 0, \end{aligned}$$

from which we can derive the expressions for emissions

$$e_1^C(t) = \alpha_1 + \mu v_S e^{\rho t}, \quad e_2^C(t) = \alpha_2 + \mu v_S e^{\rho t}.$$

The steps are the same: we choose a guess for value function, then we substitute its partial derivatives and the emissions in equation (13). So, take

$$v(t, S) = (AS + B) e^{-\rho t},$$

whose partial derivatives are: $v_t = -\rho(AS + B) e^{-\rho t}$ and $v_S = A e^{-\rho t}$. Then optimal emissions are given by

$$e_1^C(t) = \alpha_1 + \mu A, \quad e_2^C(t) = \alpha_2 + \mu A.$$

Substituting in (13), we can derive an expression for parameter A

$$\rho A = -n_1 \beta_1 - n_2 \beta_2 - A \delta,$$

so that, $A = -\frac{n_1 \beta_1 + n_2 \beta_2}{\rho + \delta}$. Then, the emission solutions for cooperative countries, for $t \in [T^C, \infty)$, are given by

$$e_i^C(t) = \alpha_i - \mu \frac{n_1 \beta_1 + n_2 \beta_2}{\rho + \delta}, \quad i = 1, 2. \quad (14)$$

Now, we can proceed to find the feedback Nash equilibrium for signatory players in the interval $[0, T^C]$. The different damage-cost function implies that the cooperative solutions solve:

$$\begin{aligned} \max_{e_1^C, e_2^C} \int_0^{T^C} & \left[n_1 \left(\alpha_1 e_1^C - \frac{1}{2} (e_1^C)^2 \right) + n_2 \left(\alpha_2 e_2^C - \frac{1}{2} (e_2^C)^2 \right) - n_1 \beta_1 S - n_2 \gamma \frac{t}{T^C} \beta_2 S + \right. \\ & \left. + s_1 n_1 + s_2 n_2 \right] e^{-\rho t} dt, \end{aligned}$$

$$\begin{aligned} \text{s.t. } \dot{S}(t) &= \mu \left(n_1 e_1^C(t) + n_2 e_2^C(t) + (N_1 - n_1) e_1^{NC}(t) + (N_2 - n_2) e_2^{NC}(t) \right) - \delta S(t), \\ S(0) &= S_0. \end{aligned}$$

If we denote with $v(t, S)$ the value function, than the HJB equation is

$$\begin{aligned} -v_t = \max_{e_1^C, e_2^C} \left\{ \left[n_1 \left(\alpha_1 e_1^C - \frac{1}{2} (e_1^C)^2 \right) + n_2 \left(\alpha_2 e_2^C - \frac{1}{2} (e_2^C)^2 \right) - n_1 \beta_1 S - n_2 \gamma \frac{t}{T^C} \beta_2 S \right] e^{-\rho t} + \right. \\ \left. + v_S \left[\mu \left(n_1 e_1^C + n_2 e_2^C + (N_1 - n_1) e_1^{NC} + (N_2 - n_2) e_2^{NC} \right) - \delta S \right] + \right. \\ \left. + (s_1 n_1 + s_2 n_2) e^{-\rho t} \right\}. \end{aligned} \quad (15)$$

As usual, we compute the first order conditions in (15), to obtain a characterization for signatories' emissions. So

$$\begin{aligned} n_1 [(\alpha_1 - e_1^C) e^{-\rho t} + v_S \mu] &= 0, \\ n_2 [(\alpha_2 - e_2^C) e^{-\rho t} + v_S \mu] &= 0, \end{aligned}$$

from which $e_i^C(t) = \alpha_i - v_S \mu$, $i = 1, 2$.

We need to give a guess for value function, and we choose, as in the previous cases, a linear function of S

$$v(t, S) = [g(t)S + z(t)]e^{-\rho t},$$

whose partial derivatives with respect to t and S are

$$v_t = [(g'(t) - \rho g(t))S + z'(t) - z(t)]e^{-\rho t}, \quad v_S = g(t)e^{-\rho t}.$$

The expression of v_S gives us the emission solutions

$$e_i^C = \alpha_i + \mu g(t), \quad i = 1, 2.$$

To conclude the computation of the Nash equilibrium, we need to find an expression for the function $g(t)$. The way is to substitute v_t , v_S and e_i , $i = 1, 2$, inside (15). With some algebra, and because of the continuity of value function, we have to solve the dynamical system

$$\begin{cases} -g'(t) + (\rho + \delta)g(t) = n_1 \beta_1 + n_2 \gamma \frac{t}{T^C} \beta_2, \\ g(T^C) = -\frac{n_1 \beta_1 + n_2 \beta_2}{\rho + \delta}. \end{cases}$$

The system has a unique solution $g(t)$, as follows

$$g(t) = -\frac{n_1\beta_1}{\rho + \delta} - \frac{n_2\beta_2}{T^C(\rho + \delta)^2} \left[\gamma(1 + t(\rho + \delta) - e^{(\rho + \delta)(t - T^C)}) + \right. \\ \left. + T^C(1 - \gamma)(\rho + \delta)e^{(\rho + \delta)(t - T^C)} \right].$$

Finally, the emissions for signatory players, when $t \in [0, T^C]$, are given by:

$$e_i^C(t) = \alpha_i - \mu \frac{n_1\beta_1}{\rho + \delta} - \mu \frac{n_2\beta_2}{T^C(\rho + \delta)^2} \left[\gamma(1 + t(\rho + \delta) - e^{(\rho + \delta)(t - T^C)}) + \right. \\ \left. + T^C(1 - \gamma)(\rho + \delta)e^{(\rho + \delta)(t - T^C)} \right],$$

where $i = 1, 2$.