

# DISCUSSION PAPER SERIES

No. 5046

**OPTIMAL TRANSFERS AND  
PARTICIPATION DECISIONS IN  
INTERNATIONAL ENVIRONMENTAL  
AGREEMENTS**

Carlo Carraro, Johan Eyckmans and  
Michael Finus

*INTERNATIONAL TRADE*



**C**entre for **E**conomic **P**olicy **R**esearch

[www.cepr.org](http://www.cepr.org)

Available online at: [www.cepr.org/pubs/dps/DP5046.asp](http://www.cepr.org/pubs/dps/DP5046.asp)

# OPTIMAL TRANSFERS AND PARTICIPATION DECISIONS IN INTERNATIONAL ENVIRONMENTAL AGREEMENTS

**Carlo Carraro**, Università di Venezia and CEPR  
**Johan Eyckmans**, Katholieke Universiteit Leuven  
**Michael Finus**, FernUniversität Hagen

Discussion Paper No. 5046  
May 2005

Centre for Economic Policy Research  
90–98 Goswell Rd, London EC1V 7RR, UK  
Tel: (44 20) 7878 2900, Fax: (44 20) 7878 2999  
Email: [cepr@cepr.org](mailto:cepr@cepr.org), Website: [www.cepr.org](http://www.cepr.org)

This Discussion Paper is issued under the auspices of the Centre's research programme in **INTERNATIONAL TRADE**. Any opinions expressed here are those of the author(s) and not those of the Centre for Economic Policy Research. Research disseminated by CEPR may include views on policy, but the Centre itself takes no institutional policy positions.

The Centre for Economic Policy Research was established in 1983 as a private educational charity, to promote independent analysis and public discussion of open economies and the relations among them. It is pluralist and non-partisan, bringing economic research to bear on the analysis of medium- and long-run policy questions. Institutional (core) finance for the Centre has been provided through major grants from the Economic and Social Research Council, under which an ESRC Resource Centre operates within CEPR; the Esmée Fairbairn Charitable Trust; and the Bank of England. These organizations do not give prior review to the Centre's publications, nor do they necessarily endorse the views expressed therein.

These Discussion Papers often represent preliminary or incomplete work, circulated to encourage discussion and comment. Citation and use of such a paper should take account of its provisional character.

Copyright: Carlo Carraro, Johan Eyckmans and Michael Finus

## ABSTRACT

### Optimal Transfers and Participation Decisions in International Environmental Agreements\*

The literature on international environmental agreements has recognized the role transfers play in encouraging participation in international environmental agreements (IEAs), but the few results achieved so far are overly specific and do not exploit the full potential of transfers for successful treaty-making. Therefore, in this paper, we develop a framework that enables us to study the role of transfers in a more systematic way. We propose a design for transfers using both internal and external financial resources and making 'welfare optimal agreements' self-enforcing. To illustrate the relevance of our transfer scheme for actual treaty-making, we use a well-known integrated assessment model of climate change to show how appropriate transfers may be able to induce almost all countries into signing a self-enforcing climate treaty.

JEL Classification: C72, H23, Q25 and Q28

Keywords: climate policy, international environmental agreements, self-enforcing international environmental agreements and transfers

Carlo Carraro  
Dipartimento di Scienze Economiche  
Università di Venezia  
San Giobbe 873  
30121 Venezia  
ITALY

Johan Eyckmans  
Centrum voor Economische Studiën  
Katholieke Universiteit Leuven  
Naamsestraat 69  
B - 3000 Leuven  
BELGIUM

Tel: (39 041) 234 9166  
Fax: (39 041) 234 9176  
Email: ccarraro@unive.it

Tel: (32 16) 32 68 44  
Fax: (32 16) 32 69 10  
Email: johan.eyckmans@econ.kuleuven.ac.be

For further Discussion Papers by this author see:  
[www.cepr.org/pubs/new-dps/dplist.asp?authorid=104121](http://www.cepr.org/pubs/new-dps/dplist.asp?authorid=104121)

For further Discussion Papers by this author see:  
[www.cepr.org/pubs/new-dps/dplist.asp?authorid=162501](http://www.cepr.org/pubs/new-dps/dplist.asp?authorid=162501)

Michael Finus  
Institute of Economic Theory  
FernUniversität in Hagen  
Profilstrasse 8  
58084 Hagen  
GERMANY

Tel: (49 2331) 987 4454  
Fax: (49 2331) 987 302  
Email: michael.finus@fernuni-hagen.de

For further Discussion Papers by this author see:  
[www.cepr.org/pubs/new-dps/dplist.asp?authorid=162502](http://www.cepr.org/pubs/new-dps/dplist.asp?authorid=162502)

\*This paper was written while M. Finus was a visiting scholar at the Katholieke Universiteit Leuven, Centrum voor Economische Studiën (K.U.Leuven-CES, Belgium). He acknowledges the financial support by the CLIMNEG 2 project funded by the Belgian Federal Science Policy Office and the kind hospitality of K.U.Leuven.

Submitted 16 April 2005

## 1. Introduction

Transfers play a prominent role in the analysis of self-enforcing international environmental agreements (IEAs). There are two reasons why this is not surprising. First, large asymmetries in the cost and benefit structure between countries may lead to a highly asymmetric distribution of the gains from cooperation that may hamper successful treaty-making. Second, IEAs provide a public good and therefore face strong free-rider incentives that might be mitigated through the use of transfers. A review of current literature (see Carraro and Siniscalco, 1998; Finus, 2001, 2003a and Tulkens, 1998) suggests that contributions in this field can be broadly divided into two categories.

The first category analyzes IEAs using the tools of cooperative coalition theory. The analysis is based on the characteristic function that assigns a worth to every coalition, which is calculated as the aggregate payoff that a coalition can secure for its members irrespective of the configuration of the players remaining outside this coalition. As well as testing whether the coalition including all players (grand coalition) is stable (in the sense of the core, for example), the main focus is on the axiomatic foundation of normatively motivated sharing schemes such as the Nash bargaining solution, the Shapley value and the Chander-Tulkens' transfer scheme. Applications are found in the context of global warming (Chander and Tulkens, 1995, 1997; Eyckmans and Tulkens, 2003; Germain et al., 1998), acid rain (Germain et al., 1996; Kaitala et al., 1995), high seas fisheries (Kaitala and Lindroos, 1998; Lindroos, 2004; Lindroos and Kaitala, 2001, Pintassilgo, 2003) and water management (Ambec and Sprumont, 2002; Lejano and Davos, 1999). This approach may be regarded as the classical cooperative method of studying coalitions. The strength of this approach lies in the generality

of results that can often be established on the basis of some standard properties, like superadditivity.<sup>5</sup>

The second category of contributions analyzes IEAs using the tools of non-cooperative coalition theory. The analysis is based on the valuation function that assigns an individual payoff to every coalition member. The value is calculated by taking into account the entire coalition structure, i.e. the partition of players inside and outside a coalition. The main focus is on explaining free-riding behavior in the context of externalities, identifying the main economic factors that determine the relative success of partial cooperation and suggesting instruments for discouraging free-riding. The advantage of the non-cooperative over the cooperative approach is that it better captures externalities between players and coalitions (Bloch 2003). However, this advantage comes at the cost of complexity, implying that results are usually less general.

In the context of the non-cooperative approach, transfers have been analyzed in their ex-ante and ex-post forms.<sup>6</sup> Ex-ante means that countries commit to a certain transfer rule before they decide upon their participation in an IEA. Ex-post means that after an agreement has formed, transfers are used to broaden an existing coalition.

The first putative paper on transfers goes back to Carraro and Siniscalco (1993). In line with intuition, they prove that transfers have no effect in the standard model with symmetric countries, given the constraint that all IEAs must be self-enforcing and that transfers must be self-financed. They proceed to analyze various forms of commitment that will enable coalitions to expand via ex-post transfers. They suggest two directions of ex-post transfers that may improve upon the status quo: 1) insiders (coalition members) “bribing” outsiders (non-coali-

---

<sup>5</sup> Roughly speaking, superadditivity means that the worth of coalitions increases with increasing participation. See section 3 for a formal definition.

<sup>6</sup> For an overview, see Finus (2003).

tion members) to join their coalition and 2) outsiders “bribing” other outsiders to join the coalition.

The idea of various forms of commitment and ex-post transfers was also pursued in later papers by Botteon and Carraro (1997), Jeppesen and Andersen (1998) and Petrakis and Xepapadeas (1996), though commitment is certainly not an assumption in line with the notion of self-interested players. Therefore, it was important to illustrate that expansions of coalitions via ex-post transfers may also be possible without commitment as this has been done by Botteon and Carraro (1997) in a simple empirical model with five heterogeneous countries. Moreover, the positive effect of ex-ante transfers was illustrated for the Nash bargaining solution and the Shapley value. A similar conclusion was confirmed by Barrett (1997) for the Shapley value using a stylized simulation model where heterogeneity was limited to two types of countries.

Later papers have looked at the effect of various ex-ante and ex-post transfers rules on the success of coalition formation and on the possibility of expanding stable coalitions (see Altamirano-Cabrera and Finus, 2004; Bosello et al., 2003, 2004; Carraro and Siniscalco, 2001; Eyckmans and Finus, 2003, 2004a; Finus et al., 2004; Weikart et al., 2004). Most of these papers used a more elaborate empirical model, looked not only at stylized transfer schemes derived from cooperative game theory, but also considered schemes that are based on various moral motives for “fair sharing”, that considered transfers via permit trading and allowed for the possibility of multiple coalitions. Roughly speaking, all papers basically confirm earlier studies in concluding that transfers can be conducive to the success of self-enforcing agreements, but that outcomes crucially depend on the particular transfer rule, the model and the data set. For instance, Barrett (2001) draws a rather pessimistic picture of the Chander-Tulkens’ transfer rule in the context of a non-cooperative coalition model, whereas Eyckmans and Finus (2003) come to a more optimistic conclusion. Moreover, Bosello et al.

(2003) find some evidence that equitable sharing rules can enhance efficiency by increasing the number of signatories of an environmental treaty, whereas Altamirano-Cabrera and Finus (2004) derive the exact opposite conclusion.

The mixed evidence and the specific results motivate us to look for a more general and rigorous approach to studying the role of transfers in the context of non-cooperative coalition theory. In particular, we want to determine the “full potential of transfers”. To this end, we go back to the roots of the analysis of IEAs, removing in this paper any unnecessary complication as a first step. That is, we assume a simple cartel formation game and apply the concept of internal and external stability. We do not consider commitments or any complication like non-transferable utility (Buchholz and Konrad 1995), monitoring and moral hazard problems (Petraakis and Xepapadeas 1996); “reputation effects” (Jeppesen and Andersen; 1998; Hoel and Schneider 1997) are also discarded in favour of the notion of optimal transfer schemes. With optimal transfers we mean transfers designed to maximize global welfare under the constraint that the underlying IEA is self-enforcing.

In what follows, we present our model in section 2. This comprises not only a theoretical part but also an empirical part in order to illustrate the usefulness and practical application of our concepts. The empirical part is based on a modified version of RICE, a well-known integrated assessment model of climate change policy (Nordhaus and Yang, 1996). In section 3, we introduce two properties of coalition formation that are suitable for analyzing all aspects of transfers in section 4. Section 5 wraps up the main findings and highlights some directions for future research.

## 2. Model

### 2.1 Theoretical Background

Coalition formation is modeled as a two-stage game. There are  $n$  players  $N = \{1, \dots, n\}$  that are countries or world regions in our empirical model and which we simplify refer to as countries in the following discussion. In the first stage, countries choose their membership: a country can either join coalition  $S$  and become a signatory or remain a singleton and non-signatory. These decisions lead to coalition structure  $C = \{S, \{\ell\}, \dots, \{n\}\}$ , i.e., a partition of players, with  $s$  signatories ( $s$  denotes the cardinality of  $S$ ) and  $n-s$  non-signatories. Given the simple structure of the first stage, a coalition structure  $C$  is fully characterized by coalition  $S$ .<sup>7</sup>

In the second stage, countries choose their economic strategies. In the context of our empirical model, economic strategies are emission abatement and capital investment (see subsection 2.2 for details). At this stage, it suffices to denote the vector of economic strategies by  $\omega(S) = (\omega_1(S), \dots, \omega_n(S))$ , given that a coalition  $S$  has formed in the first stage; we can also note that in the second stage countries receive individual payoffs  $\pi_i(\omega(S))$  that depend on the economic strategies of all countries.<sup>8</sup>

We compute the subgame-perfect equilibria of this two-stage game by backward induction. To do this, it is sufficient for strategies to constitute a Nash equilibrium at every stage. For the

---

<sup>7</sup> This simplification would not be possible if we were to allow for multiple non-trivial coalitions as for instance considered in Bosello et al. (2003, 2004), Carraro (2000), Eyckmans and Finus (2003) and Finus (2003b).

<sup>8</sup> This simple theoretical framework has often been adopted in the literature on international environmental agreements where the assumption of a coalition structure with a single coalition is the most obvious and realistic and where the game is characterized by positive externalities. A more general framework is sometimes used in coalition theory (Bloch, 2003) but would not be useful for the purpose of showing our main results.

second stage, this entails that economic strategies form a Nash equilibrium between coalition S and the n-s non-signatories.<sup>9</sup> That is:

$$\begin{aligned} \sum_{i \in S} \pi_i(\omega_S^*(S), \omega_{-S}^*(S)) &\geq \sum_{i \in S} \pi_i(\omega_S(S), \omega_{-S}^*(S)) \quad \forall \omega_S(S) \text{ and} \\ \forall i \notin S: \pi_i(\omega_S^*(S), \omega_i^*, \omega_{-i}^*(S)) &\geq \pi_i(\omega_S^*(S), \omega_i(S), \omega_{-i}^*(S)) \quad \forall \omega_i(S). \end{aligned} \quad (1)$$

where  $\omega_S(S)$  is the economic strategy vector of coalition S,  $\omega_{-S}(S)$  the vector of all other countries not belonging to S,  $\omega_i(S)$  the strategy of non-signatory i, and  $\omega_{-i}(S)$  the strategy vector of all other non-signatories except i under coalition structure S. An asterisk denotes equilibrium strategies. Computationally, this implies that non-signatories  $i \notin S$  will choose their economic strategies so as to maximize their individual payoff  $\pi_i(\omega)$ , whereas all signatories  $i \in S$  jointly maximize  $\sum_{i \in S} \pi_i(\omega)$ , the aggregate payoff of their coalition. Strategically, this means that the behaviour of non-signatories towards all other countries is selfish and non-cooperative; signatories behave cooperatively towards their fellow members (otherwise cooperation would not be worthwhile analyzing), but non-cooperatively towards outsiders. Economically, this means strategies are group (but not globally) efficient within coalition S. Hence, the equilibrium economic strategy vector  $\omega^*(S)$  corresponds to the classical “social or global optimum” if coalition S comprises all countries ( $S=N$ ), i.e. the grand coalition forms, and corresponds to the classical “Nash equilibrium” if coalition S comprises only one member ( $s=\{i\}$ ). Hence, any inefficiency, i.e., global welfare loss of coalition S compared to the global optimum stems from the fact that S is not the grand coalition.

---

<sup>9</sup> This has been called a partial agreement Nash equilibrium by Chander and Tulkens (1997). Our assumption is in line with the mainstream of the literature on coalition theory. For an overview see Bloch (2003) and Yi (2003).

Given that the second stage of the game has been solved, we define  $v_i(S) = \pi_i(\omega^*(S))$  as the valuation of country  $i$  if coalition  $S$  forms. This definition succinctly summarizes all information relevant to the second stage.

For the first stage, we define a Nash equilibrium in terms of participation.<sup>10</sup> The following two conditions must be met:

$$\text{internal stability: } v_i(S) \geq v_i(S \setminus \{i\}) \quad \forall i \in S. \quad (2)$$

$$\text{external stability: } v_i(S) \geq v_i(S \cup \{i\}) \quad \forall i \notin S. \quad (3)$$

That is, in equilibrium, no signatory belonging to coalition  $S$  has an incentive to leave its coalition in order to become a non-signatory, given the participation decisions of all other countries. By the same token, no non-signatory has an incentive to join coalition  $S$ , given the decisions of all other countries.

Regardless of whether we consider ex-ante or ex-post transfers, in a TU-framework, optimal economic strategies are not affected by transfers. Thus, valuations with transfers  $\hat{v}_i(S)$  are related to those without transfers  $v_i(S)$  simply through the relation  $\hat{v}_i(S) = v_i(S) + t_i$  where  $t_i > 0$  means receiving and  $t_i < 0$  means paying a transfer. We make the standard assumption that transfers balance, i.e.,  $\sum_{i \in N} t_i = 0$  and hence  $\sum_{i \in N} \hat{v}_i(S) = \sum_{i \in N} v_i(S)$ .<sup>11</sup> Note that in any case (with and without transfers), coalition  $S = \{i\}$  is always internally stable and coalition  $S = N$  always externally stable, which simply follows by definition.

---

<sup>10</sup> This definition of coalitional stability is due to d'Aspremont et al. (1983) and has been frequently applied in the literature on IEAs as for instance by Barrett (1994), Carraro and Siniscalco (1993), Hoel (1992) and by many scholars afterwards.

<sup>11</sup> The condition that transfers balance is equivalent to the self-financed transfer constraint in Carraro and Siniscalco (1993).

## 2.2 Empirical Background

In order to illustrate the importance of transfers for the success of coalition formation, we derive valuations from the CLIMNEG World Simulation Model (hereafter abbreviated as CWSM). CWSM is an integrated assessment, economy-climate model that extends the seminal RICE model by Nordhaus and Yang (1996).<sup>12</sup> It captures the endogenous feedback of climate change damages on production and consumption. The economic module of the CWSM consists of a dynamic, long-term, perfect foresight, Ramsey-type optimal growth model. The environmental module consists of a carbon cycle and temperature change module. The decision variables in the CWSM are investment and carbon emission reduction.

In the CWSM, the world is divided into six regions: *USA*, *JPN* (Japan), *EU* (European Union), *CHN* (China), *FSU* (Former Soviet Union) and *ROW* (Rest of the World). In every region  $i$ , and at every time  $t$ , the following budget equation describes how “potential GDP”,  $Y_{i,t}$ , can be “allocated” to consumption,  $Z_{i,t}$ , investment,  $I_{i,t}$ , emission abatement costs,  $Y_{i,t}C_i(\mu_{i,t})$ , and climate change damages,  $Y_{i,t}D_i(\Delta T_t)$ :

$$Y_{i,t} = Z_{i,t} + I_{i,t} + Y_{i,t} C_i(\mu_{i,t}) + Y_{i,t} D_i(\Delta T_t) \quad (4)$$

Output  $Y_{i,t}$  is produced with capital and labor. Capital is built up through investment and depreciates at some fixed rate. Labour supply is assumed to be inelastic. Therefore, investment  $I_{i,t}$  is the only endogenous production input and constitutes the first choice variable in the model.

Abatement costs  $Y_{i,t} C_i(\mu_{i,t})$  are expressed as “loss of potential GDP”:  $C_i$  is the share of “potential GDP” devoted to abatement, which is a function of  $\mu_{i,t} \in [0,1]$ , a variable that

---

<sup>12</sup> An overview of the equations and parameters with a detailed exposition of the model can be found in Eyckmans and Tulkens (2003).

measures the relative emission reduction compared to the business-as-usual scenario without any abatement policy. Damages  $Y_{i,t} D_i(\Delta T_t)$  are also expressed as “loss of potential GDP”:  $D_i$  is the share of “potential GDP” destroyed by climate change damages, which is a function of temperature change  $\Delta T_t$ . Temperature change depends on the stock of greenhouse gases, which in turn depends on emissions that accumulate in the atmosphere. Finally, emissions are proportional to production, but can be reduced by the abatement rate  $\mu_{i,t}$ . Hence, the second choice variable in this model is the emission abatement rate  $\mu_{i,t}$ .

Both choice variables (investment and abatement) affect output, abatement costs, damage costs and therefore also consumption, not only domestically but also abroad. This is immediately evident with regard to abatement because remaining emissions (after abatement) increases the stock of greenhouse gases, which affects environmental damages in every country. However, it is also true for investment, since capital is an input in the production process and emissions are proportional to production. Technological progress is captured by the CWSM in an exogenous fashion (the time path is taken from RICE). It increases production potential and decreases the emission-output ratio (i.e. increases energy efficiency) over time.

Welfare is measured as total lifetime discounted consumption:

$$\pi_i(\omega) = \sum_{t=0}^{\Omega} \frac{Z_{i,t}}{[1 + \rho_i]^t} \quad (5)$$

where  $\rho_i$  stands for the discount rate of region  $i$ ,  $\Omega$  denotes the time horizon and  $\omega$  is an economic strategy vector. Vector  $\omega = \{I_{i,t}, \mu_{i,t}\}_{i \in N; t=0, \dots, \Omega}$  consists of a time path of 35 decades<sup>13</sup> for emission abatement and investment<sup>13</sup> for all six regions and hence its length is  $2 \times 35 \times 6 = 420$ . For every possible coalition  $S$ , we compute the open-loop Nash equilibrium

---

<sup>13</sup> We choose a sufficiently long time period to avoid „end point bias”. However, due to discounting, only a shorter period is strategically relevant for players.

$\omega^*(S)$  in order to derive valuations  $v_i(S) = \pi_i(\omega^*(S))$  as described in subsection 2.1. Given that our empirical model comprises six players, we have 58 different coalitions and therefore a full table of valuation vectors of dimension 58x6. If valuations are modified through transfers, this happens in a one-shot fashion since it does not affect equilibrium economic strategies in the CWSM as proved in Eyckmans and Tulkens (2003). Thus, we are operating within a TU-framework.

We finish this section with five remarks about the basic incentive structure obtained by calibrating the CWSM. First, we assume a relatively low discount rate of 1.5 percent, except for CHN and ROW where we assume a discount rate of 3 percent in order not to “overestimate” the incentives for these regions to implement climate change policies. However, much higher discount rates would simply ignore the long term effects of climate change, providing no incentive for countries to cooperate and therefore would render our analysis uninteresting. The discount rates chosen are in line with the recommendations in Weitzman (2001).

Second, the parameters set for the CWSM imply that USA, JPN and EU face steep abatement cost curves, while CHN and ROW face flat ones. The regional differences in abatement costs mainly reflect differences in energy efficiency. Intuitively, energy efficient regions face higher marginal costs when cutting back emissions than regions characterized by low energy efficiency because they have already exploited the cheapest energy saving techniques.

Third, damage functions are particularly steep in EU and ROW, less steep in USA and JPN and relatively flat in FSU and CHN. The high damage estimate (as a percentage of “potential GDP”) for ROW is due to the fact that climate change is believed to affect developing countries more strongly than industrialized countries, because their economies tend to depend more on climate related production processes like agriculture, fishery and forestry (IPCC

2001). The low damage estimate for FSU is due to some expected benefits from moderate temperature increase, like the increased availability of arable land.

Fourth, in a given coalition  $S$ , the steeper the marginal damage cost curves and the flatter the marginal abatement cost curves of the members of  $S$  are, the higher the optimal abatement of coalition members will be, which follows from the first-order conditions of joint welfare maximization of coalition  $S$  (see Barrett, 1994 and Eyckmans and Tulkens, 2003). It follows that in any period, coalition members should abate up to the point where their marginal abatement costs are equal to the discounted sum of all coalition members' avoided future marginal climate change damages.

Fifth, coalition members with a flatter marginal abatement cost curve have to contribute more than those with a steeper curve, all else being equal, which also follows from the first order conditions (cost efficiency) of coalition  $S$ .

### 3. Properties of Valuations

In this section, we discuss two important properties that hold for the valuations derived from our empirical model CWSM and which determine the general incentive structure of countries in the coalition formation game. The first property is called superadditivity and means that the aggregate valuation of country  $j$  and coalition  $S$  increases if country  $j$  joins coalition  $S$ .

#### Property 1: Superadditivity

*A coalition game is superadditive if and only if for all  $S \subset N$  and  $j \notin S$ :*

$$\sum_{i \in S \cup \{j\}} v_i(S \cup \{j\}) > \sum_{i \in S} v_i(S) + v_j(S).$$

That is, there is “coalitional gain” from cooperation and hence cooperation is “group rational” or “*coalitionally rational*”. It is evident that superadditivity is a necessary condition to make

cooperation attractive for those countries participating in an IEA. This property means that starting from any coalition  $S$  and increasing the degree of cooperation by moving to  $S \cup \{j\}$  or even larger coalitions, it is generally possible to allocate the coalition gain such that it constitutes a Pareto-improvement to all regions involved in cooperation.<sup>14</sup>

We can define the second property with the term positive externalities, meaning that if country  $j$  joins coalition  $S$ , all countries that do not belong to  $S \cup \{j\}$  are better off.

### **Property 2: Positive Externalities**

*A coalition game exhibits positive externalities if and only if for all  $S \subset N$ ,  $j \notin S$  and all  $\ell \notin S \cup \{j\}$ :  $v_\ell(S \cup \{j\}) > v_\ell(S)$ .*

Consequently, there is an “external gain” or a positive spillover from cooperation, making free-riding attractive. From a non-signatory’s point of view, the most favorable condition is the one in which all other countries participate in the agreement.<sup>15</sup>

It is then clear that a region’s decision to join a coalition – as well as the stability of an IEA – depends on the intensity of the superadditivity effect which, together with the sharing rule of the coalitional gain, determines the inside options of cooperation relative to the intensity of the positive externality effect which in turn determines the outside options of cooperation. We will study these effects in more detail in section 4, but note here that superadditivity and positive externality together imply that global welfare increases through cooperation. That is, given a coalition  $S$ , whenever a single or several countries join coalition  $S$ , global welfare is

---

<sup>14</sup> Superadditivity is a property frequently encountered in cooperative coalition theory, but not much used in non-cooperative coalition theory, despite the fact that it helps to structure ideas immensely.

<sup>15</sup> Positive (and negative) externalities is a property that plays an important role in recent literature on non-cooperative coalition theory. See for instance Bloch (2003), Yi (2003) and Maskin (2003).

raised. That is, cooperation is *globally rational* - a central property that motivates our effort of analyzing measures to mitigate the problems of free-riding in transboundary pollution control.

Table 1 illustrates the magnitudes at stake for our empirical application using the CWSM model. It displays – for a selection of coalitions – welfare (global welfare) and two environmental variables (carbon concentration and global emissions) in absolute (in the legend) and relative terms (in the table). The relative magnitudes can be interpreted as a “closing the gap index”, measuring how close a coalition comes to the global optimum where the performance in the global optimum is 100 percent and the performance with no cooperation is 0 percent by definition. Apart from stressing that both full and partial cooperation can make a large difference in welfare and ecological terms compared to no cooperation, Table 1 illustrates that not only the size of a coalition matters for the global success of cooperation, but also the identity of its members. Put differently, the commonly held view that high participation automatically indicates the success of an IEA may be wrong. For instance, coalition no. 32 including five members (USA, JPN, EU, CHN, FSU) ranks lower than many coalition structures comprising coalitions of only three or four members and even lower than coalition no. 31 with only two members (JPN and ROW).

From Table 1, it is also clear that, as a general tendency, the importance for global welfare of participation of particular countries decreases with the following sequence: ROW, CHN, EU, USA, FSU and JPN. ROW’s and CHN’s important role stems from the fact that they can provide cheap abatement. Similarly, JPN’s lesser importance is due to its steep marginal abatement cost curve. However, there is also an additional dimension related to environmental damages. If a given coalition maximizes its joint welfare, the higher the marginal damages of coalition members are, the higher joint abatement efforts will be, all else being equal. This explains the importance of EU for cooperation. These remarks also explain why the “old Kyoto coalition” – comprising USA, JPN, EU and FSU – in our model ranks relatively low

since the two key players – CHN and ROW – are outsiders. A similar conclusion applies to the “present Kyoto coalition”, after the withdrawal of the USA in 2001. It is evident that the US decision implies a dramatic drop in welfare and environmental effectiveness, almost to non-cooperative levels. Thus, our model provides support for the efforts of many governments and non-governmental organization to convince the US to rejoin the Kyoto Protocol. However, it also provides support for the concern of the US and many others that an effective climate policy must include developing countries (i.e., ROW) and countries in transition (i.e., CHN).<sup>16</sup>

#### 4. Stable Coalitions

In this section, we use the CWSM model to identify stable coalitions of the climate game under the alternative assumptions of no transfers and (various forms of) transfers. To gain an understanding of the driving forces, it is important to recall that stability is defined by two components – internal and external stability – and that valuations are characterized by superadditivity as well as positive externalities. Moreover, it is helpful to note that the positive externality property implies that a necessary condition for internal stability is individual rationality. *Individual rationality*, also sometimes called profitability (see Carraro and Siniscalco, 1993), means that every coalition member  $i \in S$  receives at least the same valuation in coalition  $S$  as it does under conditions of no cooperation ( $\forall i \in S: v_i(S) \geq v_i(\{i\})$ ). In other words, a minimum requirement for coalition  $S$  to be internally stable (and therefore stable) is that cooperation should, for all members of  $S$ , constitute a (weak) Pareto-improvement compared to no cooperation. The reason is simply explained. Suppose a

---

<sup>16</sup> Similar conclusions can also be found in Buchner et al. (2002) where an integrated economy-climate model based on RICE is also used. The main difference is that in the model used by Buchner et al. (2002) technical change is endogenous.

member  $i$  in  $S$  would receive a lower valuation than under no cooperation, i.e.,  $v_i(S) < v_i(\{i\})$ . If region  $i$  were to leave coalition  $S$ , it would receive valuation  $v_i(S \setminus \{i\})$  for which  $v_i(S \setminus \{i\}) \geq v_i(\{i\})$  holds (with strict inequality if  $s \geq 3$ ) due to positive externalities. Consequently, leaving coalition  $S$  would always pay and therefore  $S$  could not be internally stable.

#### 4.1 No Transfers

In the case of no transfers, only 11 out of 58 coalitions are individually rational in our CWSM model analysis (see Table 2). None of the top 20 ranked coalitions (in terms of global welfare) is individually rational. No coalition with 5 members and only one with 4 members is individually rational (see Table 2). Neither the “old” (no. 47) nor the present (no. 50) Kyoto coalition is individually rational. Notably, not a single coalition that includes China – the key player with cheap abatement options – is individually rational. Thus, in the absence of transfers, although cooperation may be coalitionally (because of superadditivity) and globally (because of superadditivity and positive externality) rational, it may not be individually rational to all coalition members. The reason is that an efficient allocation of abatement burdens within a coalition  $S$  would result in a highly asymmetric allocation of the net gains from cooperation among coalition members that face a markedly heterogeneous benefit and cost structure. For instance, the EU has a relatively steep marginal abatement and marginal damage cost curve. Therefore, if the EU is a member of a coalition  $S$ , it will be a major beneficiary of cooperation, because it contributes relatively little to cooperation, but in proportion benefits more. The opposite is true for China, which faces a flat marginal abatement cost and damage cost curve and which therefore is a typical loser from cooperation without transfers.

Thus, even a simple check for individual rationality indicates that without transfers a key player like China cannot be convinced to join a climate treaty. Moreover, we can already

conjecture that without transfers even moderate partial cooperation will prove very difficult. This is confirmed by a detailed analysis of internal and external stability as shown in Table 2. The two individually rational coalitions with the highest global welfare (no. 21 and 22) are not internally stable, though all other individually rational coalitions are internally stable. However, none of the internally stable coalitions is also externally stable. Hence, there is no stable coalition without transfers when valuations are derived from the CWSM model (see also Table 3).

## 4.2 Ex-Ante Transfer Schemes

In the light of the above negative conclusion, we consider different ex-ante transfer schemes. That is, the membership decision in the first stage of the game is based on the assumption that coalition  $S$  will not only choose its optimal economic strategies in the second stage, but will also allocate the coalition gain from cooperation among its members with a particular transfer scheme. From section 2, we may recall that  $\hat{v}_i(S) = v_i(S) + t_i$  and that transfers balance, i.e.,  $\sum_{i \in S} t_i = 0$  and hence  $\sum_{i \in S} \hat{v}_i(S) = \sum_{i \in S} v_i(S)$ .

We start by considering three transfer schemes that have played an important role in previous analyses of self-enforcing IEAs (see the literature cited in the Introduction). We call these schemes “simple” in order to distinguish them from our “optimal” transfer schemes that we introduce subsequently. Through the illustration of both optimal and simple transfer schemes, the full potential of an optimal design of transfers will become apparent.

### 4.2.1 Simple Transfer Schemes

All three simple transfer schemes that we consider originate from cooperative coalition theory. Nevertheless, they have been frequently adopted in the context of the valuation function approach. This requires only a slight modification of their original definitions to account for the fact that coalition  $S$  may not only be the grand coalition but can be any

subcoalition of  $N$ . The following formulas describe valuations of player  $i$  being a member of a given coalition  $S \subseteq N$ .

The first transfer scheme is the *Shapley Value* and implies valuations of the following form:

$$\hat{v}_i^{SV} = \sum_{\substack{T \subseteq S \\ i \notin T}} \frac{t!(s-t-1)!}{s!} \left[ \sum_{k \in T \cup \{i\}} v_k(T \cup \{i\}) - \sum_{k \in T} v_k(T) \right] \quad \forall i \in S \quad (6)$$

with coalition  $S \subseteq N$ ,  $T \subseteq S$  a subgroup of  $S$  and  $t$  and  $s$  the size of group  $S$  and  $T$ . Roughly speaking, the *Shapley Value* gives every country a valuation according to its marginal contribution to every possible subcoalition  $T$  of  $S$  (term between square brackets in (6)), weighted by the probability that this subpartition forms (first factor in (6)).

The second simple transfer scheme is the *Nash Bargaining* solution (with equal weights):

$$\hat{v}_i^{NB} = v_i(\{i\}) + \frac{1}{s} \left[ \sum_{j \in S} v_j(S) - \sum_{j \in S} v_j(\{i\}) \right] \quad \forall i \in S \quad (7)$$

Every member in  $S$  receives its valuation under no cooperation (first term) plus an equal share of the coalitional surplus compared to no cooperation (second term). Thus, no cooperation serves as the threat point.

The third simple transfer scheme is the *Chander and Tulkens'* transfer scheme in the version as applied in Eyckmans and Tulkens (2003):

$$\hat{v}_i^{CT} = v_i(\{i\}) + \frac{D'_i}{\sum_{j \in S} D'_j} \left[ \sum_{j \in S} v_j(S) - \sum_{j \in S} v_j(\{i\}) \right] \quad \forall i \in S \quad (8)$$

with  $D'_i$  discounted marginal damages of member  $i$ . It is evident that this scheme is a version of the *Nash bargaining* rule with unequal weights. This rule gives a higher share of the gains from cooperation to those that are most affected by climate change.

It is straightforward to show that all three simple transfer schemes are all coalitionally rational, i.e.  $\forall S \subseteq N: \sum_{j \in S} \hat{v}_j(S) \geq \sum_{j \in S} v_j(S)$ , and individually rational, i.e.,  $\forall i \in S: \hat{v}_i(S) \geq v_i(\{i\})$  since superadditivity holds.

Tables 2 and 3 confirm our intuition and also many previous studies on transfers: simple transfers improve upon the outcome without transfers. However, it is important to realize that this result is by no means general. Of course, all simple transfer schemes guarantee individual rationality, but individual rationality is only a necessary condition for internal stability. Though this is not the case for our data set, in a different model it may be possible that a coalition is stable without transfers and leads to a higher global welfare than any other stable coalition with a simple transfer scheme. Already from Table 2, it can be seen that there are three coalitions of size 3 that are internally stable without transfers, but none under the Shapley Value and the Chander and Tulkens' transfer rule. Similarly, no general conclusion is possible with respect to external stability either.

Moreover, in our application, the Nash bargaining solution leads to a stable coalition (no. 16) with higher global welfare than in any other stable coalition under the other two transfer schemes. However, other models could yield different results. Finally, we have no information about whether we could do any better than the *Nash Bargaining* solution and if so what the “best” transfer scheme would be and which coalition could be achieved.

#### 4.2.2 Optimal Transfer Schemes<sup>17</sup>

In order to address the issues raised above in a systematic way, we first focus on internal stability. For this purpose, we introduce the concept of a “Potentially Internally Stable

---

<sup>17</sup> Formal proofs of all theoretical claims in this subsection can be found in Eyckmans and Finus (2004b).

Coalition” (PISC) which we define as follows (see Eyckmans and Finus, 2004b and Botteon and Carraro, 1997 for a similar concept):

**Definition 1: Potentially Internally Stable Coalition (PISC)**

*A coalition  $S$  is said to be potentially internally stable (PIS) if and only if*

$$\sum_{i \in S} v_i(S) \geq \sum_{i \in S} v_i(S \setminus \{i\}).$$

Thus, if a coalition  $S$  is not PIS, this simply means that there is no transfer scheme that can ensure internal stability to all members of  $S$  and hence this coalition cannot be stable. Conversely, if a coalition  $S$  is PIS, then this means that coalition  $S$  has sufficient resources to prevent a coalition member from leaving coalition  $S$ . Thus, what is required now is to construct a transfer scheme that ensures internal stability to all members of  $S$ , provided  $S$  is PIS. Given the design of most simple transfer schemes, it seems appropriate to construct a transfer scheme that gives every member of  $S$  its outside option  $v_i(S \setminus \{i\})$  plus a share of the coalition surplus compared to the free-rider valuation:

**Definition 2: Optimal Transfer Schemes (OPTS)**

*A transfer scheme can be termed optimal if it satisfies*

$$\forall S \subseteq N, \forall i \in S: \hat{v}_i^{OP}(S) = v_i(S \setminus \{i\}) + \lambda_i(S) \left[ \sum_{j \in S} v_j(S) - \sum_{j \in S} v_j(S \setminus \{i\}) \right]$$

$$\text{with } \lambda(S) \in \Delta^{s-1} = \left\{ \lambda \in \square_+^s \mid \sum_{j \in S} \lambda_j = 1 \right\}.$$

By the construction of OPTS, it is easy to see that any transfer scheme that belongs to the class of OPTS will make any PISC internally stable. The design of OPTS suggests that we have much leeway in choosing weights  $\lambda_i(S)$ . What is important is to choose the “correct” threat point (first term of the definition of  $\hat{v}_i^{OP}(S)$  in Definition 2) and define the coalition

surplus (second term of the definition of  $\hat{v}_i^{\text{OP}}(S)$  in Definition 2) such that coalitional rationality holds, which is the case since  $\sum_{i \in S} \hat{v}_i^{\text{OP}}(S) = \sum_{i \in S} v_i(S)$ . Note that if  $S$  is not PIS, then total coalitional payoff is insufficient to satisfy the free-riding claims of all members of  $S$  and the transfer scheme becomes a loss sharing instead of a surplus sharing formula.

It is evident that any transfer scheme that belongs to the class of OPTS will lead to the same set of internally stable coalitions. Less evident but interesting is that this “robustness result” also carries over to the set of externally stable coalitions. This is because for any OPTS, either a coalition  $S$  is internally stable for all members (if it is PIS) or fails to be internally stable for all members (if it is not PIS).

More specifically, pick a coalition  $S$  and suppose  $S$  is PIS. Then, under any OPTS,  $S$  is internally stable for all coalition members, i.e.,  $\forall i \in S: \hat{v}_i^{\text{OP}}(S) \geq v_i(S \setminus \{i\})$  by the very definition of PIS and OPTS and regardless of the particular sharing vectors  $\lambda$ .

Suppose now that  $S$  is not PIS. Then, following a similar line of argument, under OPTS for all  $i \in S$ , all coalitions  $S \setminus \{i\}$  are externally stable with respect to accession of region  $i \in S$  because coalition members receive less than their free-riding payoff  $v_i(S \setminus \{i\})$ , regardless of the particular sharing vectors  $\lambda$ . Therefore, any family of weights of an OPTS will lead to the same set of internally and externally stable coalitions.

As is clearly confirmed by Table 2, every coalition that is internally stable under a simple transfer scheme will also be internally stable under an optimal transfer scheme, but not vice versa. This is not true however for external stability. For instance, coalition no. 16 is externally stable under Nash Bargaining, but is not externally stable under an optimal transfer scheme. Hence, one may wonder whether some effort should be made to search for a transfer scheme that achieves not only internal, but also external stability in an “optimal way”. A closer inspection of the underlying fundamentals reveals that this is not necessary. First, if

coalition  $S$  is not externally stable against accession of player  $j$ , then coalition  $S \cup \{j\}$  is internally stable with respect to a withdrawal of  $j$ . This follows simply from the definition of stability (see subsection 2.1). Second, we know from above that if  $S \cup \{j\}$  is internally stable with respect to a withdrawal of  $j$ , it is also internally stable for all other members of  $S \cup \{j\}$  under OPTS. Third, due to superadditivity and positive externalities, global welfare of coalition  $S \cup \{j\}$  is higher than global welfare of coalition  $S$ . Thus, we do not need to worry about external stability from a global point of view. In particular, we can conclude that the coalition with the highest global welfare among all PISC, say  $S^*$ , is also externally stable. (If  $S^*$  were not externally stable, then there would exist a larger PISC with higher global welfare, violating the initial assumption that  $S^*$  generates the highest global welfare.)

To sum up, any transfer scheme that belongs to the class of OPTS leads to the same set of internally and externally stable coalitions. Thus, results are robust and independent of specific distributional weights. Moreover, any OPTS exploits the full potential of self-enforcing cooperation. The coalition with the highest global welfare among the potentially internal stable coalitions will be stable.

The importance of these findings are evident from Table 2 and 3. In our application, the three simple transfer schemes lead to very different equilibria. Under the simple transfer schemes, the Nash Bargaining solution generates the highest global welfare (with 68.2 percent of total maximum welfare), whereas OPTS achieves 94.5 percent of total maximum welfare for coalition no. 5 {USA, EU, CHN, ROW}, which is the coalition yielding the highest global welfare among all PIS coalitions.

We finish this subsection with two general observations. First, there always exists a stable non-trivial coalition under OPTS.<sup>18</sup> In contrast, under no transfers or simple transfer schemes, a stable non-trivial coalition may fail to exist, though this applies only to the no transfer case in our application with the CWSM model. Second, there is no general guarantee that all coalitions are individually rational under OPTS, whereas this *is* the case for all simple transfer schemes. However, this poses no problem: (i) all coalitions that are PIS are internally stable under any OPTS and therefore individually rational due to the positive externality property and (ii) coalitions that are not PIS might not be individually rational but they cannot be stabilized anyway.<sup>19</sup>

### 4.3 Ex-Post Transfer Schemes

In this subsection, we consider ex-post transfers. This means that after a coalition has formed, transfers are used to modify the status quo. The status quo is a stable coalition that has emerged from the coalition formation process based either on no transfers or on some ex-ante transfer scheme.<sup>20</sup> In our application, the status quo is represented by the stable coalitions listed in Table 2 and 3. The status quo can be modified using transfers to expand a coalition  $S$ . Following Carraro and Siniscalco (1993), we cite two cases. In the first case, a coalition  $S$  uses transfers to expand its agreement by “bribing” outsiders to join the coalition (subsection

---

<sup>18</sup> Due to superadditivity, all coalitions with two members are internally stable under OPTS. If one of them is externally stable, the claim is obvious. If none of them is externally stable, then there exist larger coalitions that are internally stable. Again, if they are externally stable, the claim is confirmed and if not, then there are even larger coalitions that are internally stable. The argument continues, noting that at least the grand coalition is externally stable by definition.

<sup>19</sup> This explains why in the column denoted by “#IR” in Table 2, no exact numbers can be given under OPTS - this would require us to assume a particular set of weights.

<sup>20</sup> It will become evident below that our arguments also apply to a wider interpretation of the status quo, which also includes non-stable coalitions.

4.3.1); in the second case, an outsider  $j \notin S$  uses transfers to bribe another outsider  $k \notin S$  to join coalition  $S$ . Again, we impose budget neutrality, meaning that transfers must balance.

#### 4.3.1 Expansion of Coalitions Through Internal Means

The standard procedure to analyze the expansion of coalitions through internal means is to pick a stable coalition (see Botteon and Carraro, 1997). For instance, in the case of the Nash Bargaining solution, one may pick coalition no. 16. Subsequently, we check to see whether expansion of coalition  $S$  is possible, where current members of  $S$  compensate an outsider  $j$  for joining their coalition. It can be argued that expansion is possible if and only if 1) the expansion constitutes a Pareto-improvement to all members of  $S$  and 2) the enlarged coalition  $S \cup \{j\}$  is internally stable. The first requirement follows from the presumption that current members of  $S$  will only bribe outsider region  $j$  to join if they are better off once it does so. The second requirement simply follows from the definition of stability.

The first requirement means that

$$\forall i \in S: \quad v_i(S \cup \{j\}) + t_i \geq v_i(S) \quad (9.a)$$

$$v_j(S \cup \{j\}) + t_j \geq v_j(S) \quad (9.b)$$

and the second requirement that

$$\forall i \in S: \quad v_i(S \cup \{j\}) + t_i \geq v_i(S \cup \{j\} \setminus \{i\}) \quad (10.a)$$

$$v_j(S \cup \{j\}) + t_j \geq v_j(S) \quad (10.b)$$

must hold where typically  $t_i \leq 0$  and  $t_j \geq 0$ . By adding (9.a) and (9.b), summing over all regions  $S \cup \{j\}$  involved in the expansion and noting that  $\sum_{i \in S \cup \{j\}} t_i = 0$ , we get

$$\sum_{i \in S \cup \{j\}} v_i(S \cup \{j\}) \geq \sum_{i \in S \cup \{j\}} v_i(S) . \quad (11)$$

Condition (11) is a necessary condition for (9) to hold and may be called Potentially Pareto-Improvement (PPI). However, by consulting Definition 1, it is evident that this condition is nothing else than the condition of superadditivity. Since we know that superadditivity holds in our global emission game, PPI is a non-binding constraint.<sup>21</sup>

A similar manipulation of the second requirement reveals that a necessary pre-requisite for condition (10) to hold is

$$\sum_{i \in S \cup \{j\}} v_i(S \cup \{j\}) \geq \sum_{i \in S \cup \{j\}} v_i(S \cup \{j\} \setminus \{i\}) \quad (12)$$

which is nothing other than the condition of potential internal stability (PIS). Hence, expansion from coalition  $S$  is possible if there exists a coalition  $S \cup \{j\}$  that is PIS.

Given these remarks, the analysis of coalition expansions is straightforward since all theoretically and empirically relevant information is already known from the previous subsection 4.2 on ex-ante transfers. More specifically, we can argue as follows.

Which coalitions qualify as potential candidates for expansion? All coalitions that are PIS and which are indicated in bold in Table 2 under OPTS. In contrast, under simple transfer schemes, not all potential candidates are known and the choice of the coalition from which expansion begins may be arbitrary.

From which of the potential candidates is expansion actually possible? From all that are not externally stable because this means that coalition  $S \cup \{j\}$  is PIS. In contrast, under a simple transfer scheme, we cannot conclude that if  $S$  is externally unstable, then  $S \cup \{j\}$  is internally or potentially internally stable. Of course, if we know that if  $S \cup \{j\}$  is internally stable,  $S \cup \{j\}$  is PIS and hence an expansion from  $S$  to  $S \cup \{j\}$  is possible. However, if  $S \cup \{j\}$  is

---

<sup>21</sup> It is evident that this is also true for any  $T \subseteq S$ .

not internally stable, we cannot conclude anything under a simple transfer scheme, so we are required to make an additional check (12), namely whether  $S \cup \{j\}$  is PIS.

Finally, and probably most importantly: should we be concerned about expansions from a global welfare point of view? Not really! The coalition with the highest global welfare among PISC is internally and externally stable as we argued in subsection 4.2. Hence, no expansion via internal means is possible from this coalition. Thus, the introduction of the concept of optimal transfer schemes renders the analysis of ex-post transfers via internal means redundant. As regards expansion via internal resources for our data set, it follows that we cannot do better than coalition no. 5.

#### 4.3.2 Expansion of Coalition by External Means

We now turn to the question of whether the expansion of a coalition  $S$  is possible through external means. This means that an outsider  $k \notin S$  “bribes” another outsider  $j \notin S$  to join coalition  $S$ . Stable coalitions for which expansions via internal means are not possible are the best, although not exclusively, potential candidates for expansions via external means.<sup>22</sup> From the previous subsection, we know that the condition of Pareto-improvement is not binding in the context of superadditivity. Therefore, we can concentrate on the condition of potential internal stability. By assumption, we know that if  $S$  is internally and externally stable, then under OPTS,  $S \cup \{j\}$  is not PIS. Consequently, expansion is only possible if and only if the enlarged coalition receives sufficient transfers to compensate for the lack of PIS and, in addition, the external player is better off despite providing these resources. That is, the conditions:

---

<sup>22</sup> Indeed, as long as expansion via internal means is possible, outsiders will benefit for free from expansion through positive spillovers, knowing that expansion is in the interest of all current members of coalition  $S$ .

$$\sum_{i \in S \cup \{j\}} v_i(S \cup \{j\}) + \sum_{i \in S \cup \{j\}} t_i \geq \sum_{i \in S \cup \{j\}} v_i(S \cup \{j\} \setminus \{i\}) \quad (13.a)$$

$$v_k(S \cup \{j\}) + t_k \geq v_k(S) \quad (13.b)$$

must hold where typically  $t_i, t_j \geq 0$  and  $t_k \leq 0$  and due to budget neutrality  $\sum_{i \in S \cup \{j\} \cup \{k\}} t_i = 0$ . That is, the positive externality effect accruing to region  $k - v_k(S \cup \{j\}) - v_k(S) \geq 0$  - must be larger than the free-riding effect -  $\sum_{i \in S \cup \{j\}} v_i(S \cup \{j\} \setminus \{i\}) - \sum_{i \in S \cup \{j\}} v_i(S \cup \{j\}) > 0$  - when coalition  $S$  is expanded to coalition  $S \cup \{j\}$ .

Figure 1 shows all possibilities of expansion from coalitions with four members that are stable in our application. Figure 1 illustrates that from coalition no. 5, which is the coalition with the highest global welfare under any ex-ante OPTS or that can be achieved via ex-post transfers using internal means, no expansion via external means is possible. However, from some other coalitions with four members, expansion that raises global welfare is possible. It is interesting to note that only the US and the EU are potential candidates that have sufficient resources to pursue a unilateral policy of “bribing” other countries to participate in a climate agreement. Given that coalition no. 5 can already be achieved without external means, it is only JPN, FSU and the US that should pursue such a policy in the interest of the global good, though among this group only the US has the means and the incentive to do so.

More generally, our results stress that - in the context of problems from free-riding - countries can play an important role, even if they do not actively participate in an IEA. By subsidizing emission abatement projects abroad, those outsiders might help the receiving country to comply with the requirements of the IEA and hence, to become a formal member of it. The results also question the standard classification of the “good” and “bad” guys in international environmental policy and open up the road for alternative efficient strategies. These results may be useful for the design of future IEAs and in particular for post-Kyoto negotiations.

In our application, a “degree of optimality” of 95.6 percent (coalition no. 4 comprising all countries except USA) can be achieved by means of a clever transfer strategy - a far more optimistic result than the one obtained from most models, though, of course, full participation cannot be achieved since expansion via external means that there remains at least one outsider.

## **5. Summary and Conclusions**

The recent literature on international environmental agreements has largely neglected the role of transfers as a tool to enhance participation in international treaties. The few existing results on transfers and international environmental agreements (IEAs) are very specific and do not exploit the full potential of transfers for successful treaty-making.

In this paper, we have proposed a transfer mechanism that is capable of maximizing global welfare provided the underlying IEA is self-enforcing. This result can be achieved by using either ex-ante or ex-post transfers where in the latter case we distinguished between internal and external transfers. We have also analyzed the relationships between different types of transfers and provided a comprehensive analytical framework to address the issue of transfer design in international agreements. For example, we have shown that ex-post internal transfers are redundant if ex-ante transfers are appropriately designed. By contrast, ex-post external transfers can help to achieve a welfare improving, self enforcing IEA (compared to a situation without ex-post transfers).

To show the relevance of the theoretical results and the feasibility of the transfer scheme in actual negotiations, we used a well-known integrated assessment model of climate change policy to identify the agreement that would emerge at the equilibrium and how the transfer scheme may be able to induce all or almost all countries to sign a climate treaty. Our empirical results highlight the difficulty of reaching an effective agreement on climate policy,

because of the large asymmetries across countries and of the incentives to free-ride on the provision of a public good like climate change control. However, we also showed that the situation can be largely improved by adopting an appropriate transfer scheme. After transfers, all countries are better off, global welfare is close to the optimum and emissions are drastically reduced. We highlighted that transfers may be paid both by signatories and by non-signatories. In the former case, we proposed an “optimal” transfer scheme that leads to far better results than the “simple” transfer schemes well-known in existing literature. In the latter case, we demonstrated that it is possible to reach a self enforcing agreement involving all but one country (e.g., USA in our empirical climate model) but where this outsider contributes financially, in its self interest, to stabilize a successful agreement.

The latter result has important implications for post-Kyoto climate policy negotiations. They highlight the important role that outsiders like the USA can play. Although they did not ratify the Kyoto Protocol and may not even join this agreement at a later stage, they might nevertheless be involved in future protocols via new financial mechanisms through which they could sponsor developing countries to assume quantitative emission ceilings and to become a full member of the Protocol. These transfers may well be in the interest of the USA if the resulting increase in global greenhouse gas emission control and ensuing reduction of climate change damages outweighs the cost of the transfers and the cost of taking domestic emission control measures.

Although our numerical results are obtained using a specific model of the global economy, in the paper we have also identified the theoretical mechanisms which explain our main conclusions. These results depend only on very general properties of the underlying valuation functions, in particular superadditivity and positive externalities. Therefore, we could claim that our theoretical results are likely to hold for most models of transboundary pollution control (indeed, this is the case for the model used in Bosello et al. 2003) or even for other

types of international public good provision problems (like for instance disease control or fighting international terrorism). Of course, the particular coalition structures that can be made self enforcing will always depend on the parameterization of the simulation models.

The analysis of this paper could be extended both from a theoretical and an empirical viewpoint. A theoretical analysis would be necessary to analyze the role of transfers for example when multiple coalitions can form or when uncertainty characterizes the coalition formation game. The relationship between the optimal transfer scheme proposed in this paper and the transfers scheme implicit in the implementation of an emission trading market is also worth additional research (see Bosello et al. 2004 for some preliminary numerical results). An empirical analysis would be useful to test the robustness of our policy conclusions to other model specifications, different parameterizations or regional disaggregation.

## References

- Altamirano-Cabrera, J.-C. and M. Finus (2004), Permit Trading and Stability of International Climate Agreements. Working Papers in Economics, No. 367, University of Hagen, Germany. Forthcoming *Journal of Applied Economics*.
- Ambec, S. and Y. Sprumont (2002), Sharing a River. *Journal of Economic Theory*, vol. 107, pp. 453-462.
- Barrett, S. (1994), Self-Enforcing International Environmental Agreements. *Oxford Economic Papers*, vol. 46, pp. 804-878.
- Barrett, S. (1997), Heterogeneous International Agreements. In: C. Carraro (ed.), *International Environmental Negotiations: Strategic Policy Issues*. Edward Elgar, Cheltenham, pp. 9-25.
- Barrett, S. (2001), International Cooperation for Sale. *European Economic Review*, vol. 45 (10), pp. 1835-1850.
- Bloch, F. (2003), Non-Cooperative Models of Coalition Formation in Games with Spillovers. In: Carraro, C. (ed.), *Endogenous Formation of Economic Coalitions*, Edward Elgar, Cheltenham, UK, ch. 2, pp. 35-79.
- Bosello, F., B. Buchner and C. Carraro (2003), Equity, Development, and Climate Change Control. *Journal of the European Economic Association*, vol. 1 (2-3), pp. 601-611.
- Bosello, F., B. Buchner, C. Carraro and D. Raggi (2004), Can Equity Enhance Efficiency? Some Lessons from Climate Negotiations. In: C. Carraro and V. Fragnelli (eds.), *Game Practice and the Environment*, E. Elgar, 2004.
- Botteon, M. and C. Carraro (1997), Burden-Sharing and Coalition Stability in Environmental Negotiations with Asymmetric Countries. In: Carraro, C. (ed.), *International Environmental Negotiations: Strategic Policy Issues*. Edward Elgar, Cheltenham et al., ch. 3, pp. 26-55.
- Buchholz, W. and K. A. Konrad (1995), Strategic Transfers and Private Provision of Public Goods. *Journal of Public Economics*, vol. 57, pp. 489-505.
- Buchner, B, Carraro, C. and I Cersosimo (2002), Economic Consequences of the US Withdrawal from the Kyoto Protocol, *Climate Policy*, 76(2002), 1-20.
- Carraro, C. (2000), Roads towards International Environmental Agreements. In Siebert, H. (ed.), *The Economics of International Environmental Problems*, Mohr Siebeck, Tübingen, pp. 169-202.
- Carraro, C. and D. Siniscalco (1993), Strategies for the International Protection of the Environment. *Journal of Public Economics*, vol. 52, pp. 309-328.
- Carraro, C. and D. Siniscalco (1998), International Environmental Agreements: Incentives and Political Economy”, *European Economic Review*, 42, 561-572.

- Carraro, C. and D. Siniscalco (2001), Transfers, Commitments and Issue Linkage in International Environmental Negotiations. In: A. Ulph (ed.) *Environmental Policy, International Agreements and International Trade*, Oxford, Oxford University Press.
- Chander, P. and H. Tulkens (1995), A Core-Theoretic Solution for the Design of Cooperative Agreements on Transfrontier Pollution. *International Tax and Public Finance*, vol. 2, pp. 279-293.
- Chander, P. and H. Tulkens (1997), The Core of an Economy with Multilateral Environmental Externalities. *International Journal of Game Theory*, vol. 26, pp. 379-401.
- d'Aspremont, C., A. Jacquemin, J. J. Gabszewicz and J. A. Weymark (1983), On the Stability of Collusive Price Leadership. *Canadian Journal of Economics*, vol. 16 (1), pp. 17-25.
- Eyckmans, J. and Finus, M. (2004a), An Empirical Assessment of Measures to Enhance the Success of Global Climate Treaties (CLIMNEG Working Paper No. 61, University of Leuven (K.U.L.), Belgium).
- Eyckmans, J. and Finus, M. (2004b), An Almost Ideal Sharing Scheme for Coalition Games with Externalities. CLIMNEG Working Paper No. 62, University of Leuven (K.U.L.), Belgium.
- Eyckmans, J., and Finus, M. (2003), Coalition Formation in a Global Warming Game: How the Design of Protocols Affects the Success of Environmental Treaty-Making (CLIMNEG Working Paper No. 56, University of Leuven (K.U.L.), Belgium).
- Eyckmans, J. and H. Tulkens (2003), Simulating Coalitionally Stable Burden Sharing Agreements for the Climate Change Problem. *Resource and Energy Economics*, vol. 25, pp. 299-327.
- Finus, M. (2001), *Game Theory and International Environmental Cooperation*. Edward Elgar, Cheltenham.
- Finus, M. (2003a), Stability and Design of International Environmental Agreements: The Case of Transboundary Pollution. In: Folmer, H. and T. Tietenberg (eds.), *International Yearbook of Environmental and Resource Economics*, 2003/4, Edward Elgar, Cheltenham, UK, ch. 3, pp. 82-158.
- Finus, M. (2003b), New Developments in Coalition Theory: An Application to the Case of Global Pollution. In: Marsiliani, L., M. Rauscher and C. Withagen (eds.), *Environmental Policy in an International Perspective*, Kluwer, Dordrecht, Holland, pp. 19-49.
- Finus, M. E. Sáiz and E. M.T. Hendrix (2004), An Empirical Test of New Developments in Coalition Theory for the Design of International Environmental Agreements. Mansholt Working Paper Series, No. 14, University of Wageningen, The Netherlands.
- Germain, M., P. L. Toint and H. Tulkens (1996), International Negotiations on Acid Rains in Northern Europe: A Discrete Time Iterative Process. In: Xepapadeas, A. (ed.), *Economic Policy for the Environment and Natural Resources*, Edward Elgar, Cheltenham, UK, ch. 10, 217-236.

- Germain, M., P. L. Toint and H. Tulkens (1998), Financial Transfers to Sustain Cooperative International Optimality in Stock Pollutant Abatement. In: Faucheux, S., J. Gowdy and I. Nicolai (eds), *Sustainability and Firms: Technological Change and the Changing Regulatory Environment*. Edward Elgar, Cheltenham, UK, ch. 11, pp. 205-219.
- Hoel, M. (1992), International Environment Conventions: The Case of Uniform Reductions of Emissions. *Environmental and Resource Economics*, vol. 2, pp. 141-159.
- Hoel, M. and K. Schneider (1997), Incentives to Participate in an International Environmental Agreement. *Environmental and Resource Economics*, vol. 9, pp. 153-170.
- IPCC (2001), Climate Change 2001: Mitigation. Contribution of Working Group III to the Third Assessment Report of the Intergovernmental Panel on Climate Change, Cambridge: Cambridge University Press.
- Jeppesen, T. and P. Andersen (1998), Commitment and Fairness in Environmental Games. In: Hanley, N. and H. Folmer (eds.), *Game Theory and the Environment*. ch. 4, pp. 65-83, Edward Elgar, Cheltenham et al.
- Kaitala, V. and M. Lindroos (1998), Sharing the Benefits of Cooperation in High Seas Fisheries: a Characteristic Function Game Approach. *Natural Resource Modeling*, vol. 11, pp. 275-299.
- Kaitala, V., K.-G. Mäler and H. Tulkens (1995), The Acid Rain Game as a Resource Allocation Process with an Application to the International Cooperation among Finland, Russia and Estonia. *Scandinavian Journal of Economics*, vol. 97, pp. 325-343.
- Lejano, R. P. and C. A. Davos (1999), Cooperative Solutions for Sustainable Resource Management. *Environmental Management*, vol. 24, pp. 167-175.
- Lindroos, M. (2004), Sharing the Benefits of Cooperation in the Norwegian Spring-Spawning Herring Fishery. *International Game Theory Review*, vol. 6, pp. 35-53.
- Lindroos, M. and V. Kaitala (2001), Nash Equilibria in a Coalition Game of the Norwegian Spring-spawning Herring Fishery. *Marine Resource Economics*, vol. 15, pp. 321-339.
- Maskin, E. (2003), Bargaining, Coalitions, and Externalities. Mimeo, paper presented at the EEA-ESEM Meeting in Stockholm, August 20-24, 2003.
- Nordhaus, W.D. and Z. Yang (1996), A Regional Dynamic General-equilibrium Model of Alternative Climate-change Strategies. *American Economic Review*, vol. 86, pp. 741-765.
- Petrakis, E. and A. Xepapadeas (1996), Environmental Consciousness and Moral Hazard in International Agreements to Protect the Environment. *Journal of Public Economics*, vol. 60, pp. 95-110.
- Pintassilgo, P. (2003), A Coalition Approach to the Management of High Seas Fisheries in the Presence of Externalities. *Natural Resource Modeling*, vol. 16 (2), pp. 175-197.
- Tulkens, H. (1998), Cooperation versus Free-Riding in International Environmental Affairs: Two Approaches. In: Hanley, N. and H. Folmer (eds.), *Game Theory and the Environment*. Edward Elgar, Cheltenham et al., ch. 2, pp. 30-44.

- Weikart, H.-P., M. Finus, J.-C. Altamirano-Cabrera (2004), The Impact of Surplus Sharing on the Stability of International Climate Coalitions. Fondazione Eni Enrico Mattei Working Paper Series, No.99.04, Milano, Italy.
- Weitzman, M.L. (2001), Gamma discounting. *American Economic Review*, vol. 91 (1), pp. 260-271.
- Yi, S.-S. (2003), Endogenous Formation of Economic Coalitions: A Survey of the Partition Function Approach. In: Carraro, C. (ed.), *Endogenous Formation of Economic Coalitions*, Edward Elgar, Cheltenham, UK, ch. 3, pp. 80-127.

Table 1: Welfare and Ecological Implications of Different Coalitions\*

Coalition			Welfare	Concentration	Cumulative Emissions
N°	Size	Membership			
1	6	Grand Coalition (Full Cooperation)	100.0	100.0	100.0
2	5	USA, EU, CHN, FSU, ROW	99.1	92.2	93.0
3	5	USA, JPN, EU, CHN, ROW	96.6	90.0	91.1
4	5	JPN, EU, CHN, FSU, ROW	95.6	80.6	81.9
5	4	USA, EU, CHN, ROW	94.5	82.0	83.2
6	5	USA, JPN, CHN, FSU, ROW	93.2	73.2	74.8
7	4	EU, CHN, FSU, ROW	91.3	72.3	73.6
8	4	JPN, EU, CHN, ROW	89.6	69.8	71.5
9	4	USA, CHN, FSU, ROW	87.4	64.1	65.7
10	4	USA, JPN, CHN, ROW	85.9	61.8	63.9
11	3	EU, CHN, ROW	84.0	60.7	62.6
12	3	USA, CHN, ROW	78.8	52.0	54.3
13	4	JPN, CHN, FSU, ROW	78.4	50.3	52.6
14	5	USA, JPN, EU, FSU, ROW	70.3	66.0	67.1
15	4	USA, EU, FSU, ROW	69.1	61.0	62.0
	...	...	...	...	...
31	2	JPN, ROW	46.4	24.7	26.8
32	5	USA, JPN, EU, CHN, FSU	31.0	26.9	27.5
33	4	USA, EU, CHN, FSU	29.0	24.5	25.0
...	...	...	...	...	...
47	4	USA, JPN, EU, FSU („old Kyoto“)	5.07	1.58	2.14
...	...	...	...	...	...
50	3	JPN, EU, FSU (“present Kyoto”)	2.9	0.7	1.0
...	...	...	...	...	...
57	2	JPN, EU	0.6	0.2	0.3
58	1	Only Singleton Coalitions (No Cooperation)	0.0	0.0	0.0

\* *N°*: coalition number according to welfare ranking. *Size*: number of coalition members. *Membership*: composition of coalition. *Welfare*: global welfare expressed in relative terms:  $(\sum_{i=1}^N (w_i(c^P) - w_i(c^N))) / (\sum_{i=1}^N (w_i(c^F) - w_i(c^N)))$  where welfare is discounted lifetime consumption integrated over 1990-2300, global welfare with full cooperation is  $\sum_{i=1}^N w_i(c^F) = 339,831$  bill US\$<sub>1990</sub> (billion US dollars expressed in 1990 levels), global welfare with no cooperation is  $\sum_{i=1}^N w_i(c^N) = 338,060$  bill US\$<sub>1990</sub> and global welfare with partial cooperation is denoted by  $\sum_{i=1}^N w_i(c^P)$ . *Concentration*: atmospheric carbon concentration *M* at time *t*=2300 expressed in relative terms:  $(M(c^N) - M(c^P)) / (M(c^N) - M(c^F))$  where concentration with full cooperation is  $M(c^F) = 1912.907$  GtC (giga tons carbon), concentration with no cooperation is  $M(c^N) = 4550.202$  GtC and concentration with partial cooperation is denoted by  $M(c^P)$ . *Cumulative Emissions*: cumulative emissions of carbon over time interval 1990-2300 expressed in relative terms:  $(CE(c^N) - CE(c^P)) / (CE(c^N) - CE(c^F))$  where cumulative emissions with full cooperation are  $CE(c^F) = 772.529$  GtC, cumulative emissions with no cooperation are  $CE(c^N) = 1593.398$  GtC and cumulative emissions with partial cooperation are denoted by  $CE(c^P)$ .

**Table 2: Stable Coalition Structures Under Different Transfer Schemes**

Size	Coalitions	#S	#IR	#IS	#ES
<b>No Transfers</b>					
6	<i>1</i>	1	0	0	1
5	<i>2,3,4,6,14,32</i>	6	0	0	3
4	<i>5,7,8,9,10,13,15,18,19,21,33,34,35,37,47</i>	15	1	0	3
3	<i>11,12,16,17,20,22,23,24,25,28,36,38,39,40,41,44,48,49,50,51</i>	20	4	3	3
2	<b>26,27,29,30,31,42,43,45,46,52,53,54,55,56,57</b>	15	5	5	2
1	<b>58</b>	1	1	1	0
<b>Shapley Value</b>					
6	<i>1</i>	1	1	0	1
5	<i>2,3,4,6,14,32</i>	6	6	0	4
4	<i>5,7,8,9,10,13,15,18,19,21,33,34,35,37,47</i>	15	15	0	6
3	<i>11,12,16,17,20,22,23,24,25,28,36,38,39,40,41,44,48,49,50,51</i>	20	20	0	4
2	<b>26,27,29,30,31,42,43,45,46,52,53,54,55,56,57</b>	15	15	15	1
1	<b>58</b>	1	1	1	0
<b>Nash Bargaining</b>					
6	<i>1</i>	1	1	0	1
5	<i>2,3,4,6,14,32</i>	6	6	0	3
4	<i>5,7,8,9,10,13,15,18,19,21,33,34,35,37,47</i>	15	15	0	3
3	<i>11,12,16,17,20,22,23,24,25,28,36,38,39,40,41,44,48,49,50,51</i>	20	20	2	2
2	<b>26,27,29,30,31,42,43,45,46,52,53,54,55,56,57</b>	15	15	15	0
1	<b>58</b>	1	1	1	0
<b>Chander-Tulkens</b>					
6	<i>1</i>	1	1	0	1
5	<i>2,3,4,6,14,32</i>	6	6	0	6
4	<i>5,7,8,9,10,13,15,18,19,21,33,34,35,37,47</i>	15	15	0	10
3	<i>11,12,16,17,20,22,23,24,25,28,36,38,39,40,41,44,48,49,50,51</i>	20	20	0	10
2	<b>26,27,29,30,31,42,43,45,46,52,53,54,55,56,57</b>	15	15	15	5
1	<b>58</b>	1	1	1	0
<b>AITS</b>					
6	<i>1</i>	1	?	0	1
5	<i>2,3,4,6,14,32</i>	6	?	0	6
4	<i>5,7,8,9,10,13,15,18,19,21,33,34,35,37,47</i>	15	?	10	15
3	<i>11,12,16,17,20,22,23,24,25,28,36,38,39,40,41,44,48,49,50,51</i>	20	?	18	0
2	<b>26,27,29,30,31,42,43,45,46,52,53,54,55,56,57</b>	15	?	15	0
1	<b>58</b>	1	?	1	0

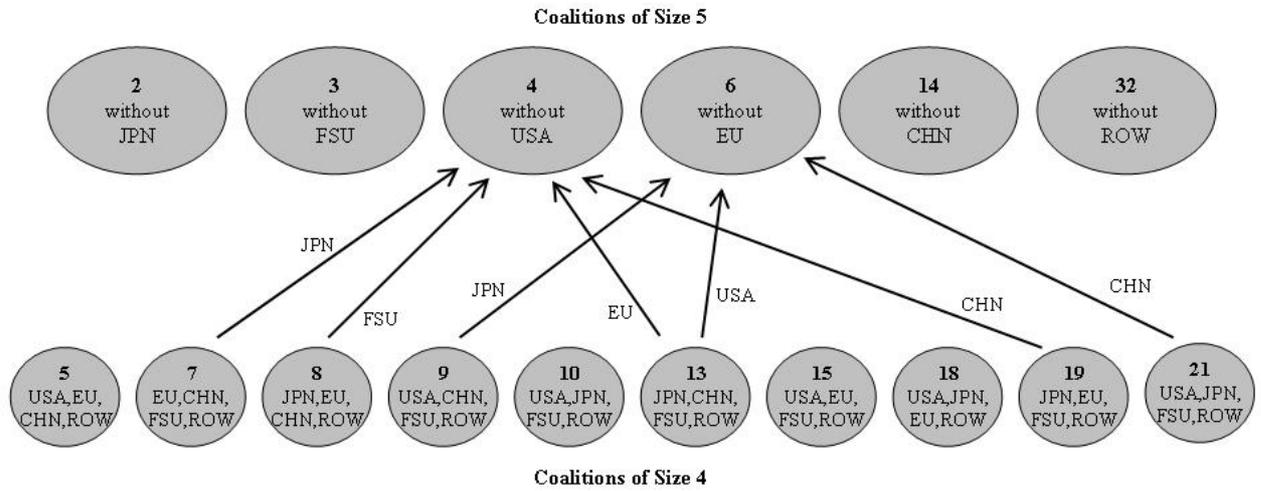
**Bold** faced means internally stable, underlined means externally stable, *italic* means NOT Individually Rational coalitions. Numbers refer to ranking according to global welfare. #S: number of coalitions of particular size, #IR: number of Individually Rational coalitions, #IS: number of Internally Stable coalitions, #ES: number of Externally Stable coalitions. For the AITS scenario, it is impossible to determine Individual Rationality without knowing the weights of the particular AITS transfer scheme.

**Table 3: Stable Coalition Structures Under Different Transfer Schemes\***

	N°	Membership	Size	Welfare	Concentration	Cumulative Emissions
	1	Grand Coalition (Full Cooperation)	6	100.00	100.00	100.00
No Transfers	-	-	-	-	-	-
Shapley	27	CHN,ROW	2	54.57	21.36	25.19
Nash Bargaining	16	CHN,FSU,ROW	3	68.21	39.15	41.63
Chander-Tulkens	26	EU,ROW	2	57.42	40.79	42.11
	27	CHN,ROW	2	54.57	21.36	25.19
	29	USA,ROW	2	54.07	35.18	36.70
	30	FSU,ROW	2	47.18	26.75	28.41
	31	JPN,ROW	2	46.39	24.72	26.80
AITS	5	USA,EU,CHN,ROW	4	94.50	81.96	83.18
	7	EU,CHN,FSU,ROW	4	91.17	72.26	73.61
	8	JPN,EU,CHN,ROW	4	89.41	69.75	71.53
	9	USA,CHN,FSU,ROW	4	87.31	64.08	65.75
	10	USA,JPN,CHN,ROW	4	85.99	61.80	63.91
	13	JPN,CHN,FSU,ROW	4	78.28	50.29	52.58
	15	USA,EU,FSU,ROW	4	68.96	61.02	62.01
	18	USA,JPN,EU,ROW	4	66.80	59.47	60.54
	19	JPN,EU,FSU,ROW	4	66.12	53.62	54.78
	21	USA,JPN,FSU,ROW	4	64.67	48.90	50.21
	58	Only Singleton Coalitions (No Cooperation)	1	0.00	0.00	0.00

\* The same legend as with Table 1 applies.

Figure 1: Enlarging Participation through External Means



**Legend:**

Bold numbers refer to the ranking of coalitions in terms of global welfare;  
description of coalition at level 5 means a coalition of 5 members, i.e., full participation without region j;  
description of coalition at level 4, the description includes the 4 coalition members that are listed;  
arrow means enlargement through bribing is possible from coalition at level 4 to coalition at level 5;  
region i attached to an arrow is the region that can be bribed to join coalition at level 4 to form coalition at level 5  
through region j listed under coalition structure at level 5 (i.e., "without j").