D
Dipartimento

S
Scienze

E
Economiche

Working Paper

Department
of Economics

Ca' Foscari University of
Venice

Carlo Carraro
Johan Eyckmans
Michael Finus

Optimal Transfers and
Participation Decisions in
International Environmental
Agreements

Optimal Transfers and Participation Decisions in International Environmental Agreements

Carlo Carraro

Corresponding author. University of Venice and Fondazione Eni Enrico Mattei FEEM

Johan Eyckmans

*European University College Brussels EHSAL and Center for Economic Studies,
Katholieke Universiteit Leuven*

Michael Finus

Department of Economics, University of Hagen and National University of Singapore

Revised Version July 2006

Abstract

The literature on international environmental agreements has recognized the role transfers play in encouraging participation in international environmental agreements. However, the results achieved so far are overly specific. Therefore, we develop a more general framework that enables us to study the role of transfers in a systematic way. We propose transfers using both internal and external financial resources for making “welfare optimal agreements” self-enforcing. To illustrate the relevance of our transfer scheme, we use a stylized integrated assessment simulation model of climate change to show how appropriate transfers may induce almost all countries into signing a self-enforcing climate treaty

Keywords

Self-enforcing International Environmental Agreements, Climate Policy, Transfers

JEL Codes

C72, H23, Q25, Q28

Address for correspondence:

Carlo Carraro
Department of Economics
Ca' Foscari University of Venice
Cannaregio 873, Fondamenta S.Giobbe
30121 Venezia - Italy
Phone: (+39) 041 2349166
Fax: (+39) 041 2349176
e-mail: ccarraro@unive.it

This Working Paper is published under the auspices of the Department of Economics of the Ca' Foscari University of Venice. Opinions expressed herein are those of the authors and not those of the Department. The Working Paper series is designed to divulge preliminary or incomplete work, circulated to favour discussion and comments. Citation of this paper should consider its provisional character.

1. Introduction

Transfers play a prominent role in the analysis of self-enforcing international environmental agreements (IEAs) (e.g., Carraro and Siniscalco 1998 and Finus 2001 and 2003a). 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. Both of these sources of potential instability of IEAs might be mitigated through the use of appropriate transfers.

The majority of recent contributions to the literature on IEAs uses the tools of non-cooperative coalition theory (see Bloch 2003 for an overview). The analysis is based on the valuation function that assigns an individual payoff to every coalition member 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.

In the context of the non-cooperative approach, transfers have been analyzed in their ex-ante and ex-post forms (see Finus 2003a for an overview). 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). They analyze various forms of commitment that enables coalitions to expand via ex-post transfers and they suggest two types of transfers that may improve upon the status quo: 1) insiders (coalition members) “bribing” outsiders (non-coalition 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.

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 (e.g., Altamirano-Cabrera and Finus 2006, Bosello et al. 2003, 2004, Carraro and Siniscalco 2001, Eyckmans and Finus 2003, 2004a, Finus et al. 2004 and Weikard et al. 2006). 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”, 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 details of the model and the data set.

The mixed evidence and the specific nature of results motivate us to look for a more general approach to study 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 employ a notion of “optimal” transfer schemes meaning transfers designed to maximize global welfare under the constraint that the underlying IEA is self-enforcing.

The importance of this research issue for policy making is evident when considering for instance the crucial role financial or technology transfers are playing in recent discussions on possible post-Kyoto climate policies in the framework of the United Nations Framework Convention on Climate Change (UNFCCC). First, multilateral financial transfers take place under the so-called Global Environment Facility (GEF) program. Though not large in absolute terms, GEF allocates and disburses currently about 250 million dollars per year in projects for improving energy efficiency, fostering renewable energy use, and pushing for sustainable transportation. Second, the flexible mechanisms (i.e. Emissions Trading ET, Joint Implementation JI and Clean Development Mechanism CDM) foreseen in the Kyoto Protocol are expected to generate important flows of financial resources among the Protocol member states and even between signatories and non-signatories in the case of CDM projects. These financial and flexibility mechanisms can be interpreted as multilateral transfer schemes. Therefore, it is important to get new insights on how transfer schemes should be designed in order to achieve maximum success of future, voluntary international climate policy agreements.

In what follows, we present our model in section 2. This comprises not only a theoretical but also an empirical part. The empirical part is based on a modified version of RICE (Nordhaus and Yang 1996) - a well-known integrated assessment model of climate change policy – which is used to illustrate our concepts. In section 3, we analyze the impact of different transfer designs on the formation of stable coalitions. Section 4 concludes.

2. Model

2.1 Theoretical Background

Coalition formation is modeled as a two-stage game. There are n players and $N = \{1, \dots, n\}$ denotes the set of countries or regions which we simplify refer to as countries in the sequel. In

the first stage, countries choose their membership: a country can either join coalition $S \subseteq N$ and become a signatory or remain a singleton and non-signatory. These decisions lead to coalition structure $C = \{S, \{1\}, \dots, \{n\}\}$, 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 .¹

In the second stage, countries choose their economic strategies. Economic strategies include for instance policy variables like emission abatement and capital investment as described in subsection 2.2. At this stage, it suffices to denote the vector of economic strategies by $\omega(S) = (\omega_1(S), \dots, \omega_n(S))$ with associated individual payoffs $\pi_i(\omega(S))$, given that a coalition S has formed in the first stage.

We compute the subgame-perfect equilibria of this two-stage game by backward induction.² To do this, it is sufficient for strategies to constitute a Nash equilibrium at every stage. For the second stage, this entails that economic strategies form a Nash equilibrium between coalition S and the $n-s$ non-signatories. This implies that non-signatories $i \notin S$ choose their economic strategies so as to maximize their individual payoff $\pi_i(\omega)$, whereas all signatories $i \in S$ jointly maximize the aggregate payoff of their coalition, $\sum_{i \in S} \pi_i(\omega)$. Strategically, this means that the behavior of non-signatories towards all other countries is selfish and non-cooperative; signatories behave cooperatively towards their fellow members, but non-cooperatively towards outsiders. Economically, this means strategies are efficient within coalition S (but not globally). Hence, the equilibrium economic strategy vector $\omega^*(S)$ corresponds to the “social optimum” if coalition S comprises all countries ($S=N$), i.e. the grand coalition forms, and

¹ The assumption of a single coalition seems to be the most obvious and realistic in the context of IEAs. Therefore, it has been frequently adopted in the literature. Exceptions are considered in Bosello et al. (2003, 2004), Carraro (2000), Eyckmans and Finus (2006) and Finus (2003b).

² Our assumption is in line with the mainstream of the literature on coalition theory. For an overview see Bloch (2003) and Yi (2003).

corresponds to the “Nash equilibrium” if coalition S comprises only one member ($S=\{i\}$). Hence, any inefficiency, i.e., global welfare loss compared to the global Pareto optimum, stems from the fact that S is not the grand coalition.

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 summarizes all information relevant to the second stage.

For the first stage, we define a Nash equilibrium in terms of participation:³

$$\text{internal stability:} \quad v_i(S) \geq v_i(S \setminus \{i\}) \quad \forall i \in S. \quad (1)$$

$$\text{external stability:} \quad v_i(S) \geq v_i(S \cup \{i\}) \quad \forall i \notin S. \quad (2)$$

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. Note that by definition coalition $S=\{i\}$ is internally stable and coalition $S=N$ externally stable by definition.

Regardless of whether we consider ex-ante or ex-post transfers, we assume that optimal economic strategies are not affected by transfers.⁴ 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

³ This definition 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.

⁴ That is, we follow the standard assumption of a transferable utility (TU) framework in the literature on coalitions for simplicity. This means to neglect income effects induced by transfers. For our simulation model this seems a reasonable assumption, given the relative small size of transfers compared to total GDP.

transfers balance, i.e., $\sum_{i \in N} t_i = 0$ and hence $\sum_{i \in N} \hat{v}_i(S) = \sum_{i \in N} v_i(S)$ (e.g., Carraro and Siniscalco 1993).

2.2 Simulation Model

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 captures the endogenous feedback of climate change damages on production and consumption.⁵ The strategic policy 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, $Y_{i,t} D_i(\Delta T_t)$:

$$Y_{i,t} = Z_{i,t} + I_{i,t} + Y_{i,t} C_i(\mu_{i,t}) + Y_{i,t} D_i(\Delta T_t) \quad (3)$$

Output $Y_{i,t}$ is produced with capital and labor. Capital is built up through investment which is the first strategic policy variable in the model. Labour supply is assumed to be inelastic. 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 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”:

⁵ Almost all parameter values are taken from the RICE model of Nordhaus and Yang (1996). A complete overview of the equations and parameter values of the CWS model can be found in Eyckmans and Tulkens (2003). The CWS model code and GAMS program to check coalition stability is available at <http://www.climneg.be>.

D_i is the share of “potential GDP” destroyed by climate change damages, which is a function of temperature change ΔT_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 $\mu_{i,t}$ which constitutes the second strategic choice variable in the model.

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.

For every of the 58 possible coalitions S , we compute the Nash equilibrium $\omega^*(S)$ between S and $N \setminus S$ in order to derive valuations $v_i(S) = \pi_i(\omega^*(S))$. That is, we determine investment and abatement strategies that maximize welfare (i.e. lifetime discounted consumption) of coalition S and of singletons in $N \setminus S$ for given investment and abatement strategies of the other players as described in subsection 2.1.⁶

2.3 Basic Incentive Structure

In this section, we briefly discuss the basic incentive structure implied by the calibration of CWSM. First, USA, JPN and EU face steep abatement cost curves, while CHN, FSU and ROW face flat ones. The regional differences in abatement costs mainly reflect differences in energy efficiency. Energy efficient regions face higher marginal costs when cutting back

⁶ This implicitly assumes that some harmonized carbon tax or a tradable permit system is implemented within coalition S such that every polluting firm in S is confronted with the coalitional damage costs of its emission behaviour. Thus, the policy instruments coordinates private investment and emission reduction decisions within coalition S .

emissions than regions characterized by low energy efficiency because they have already exploited the cheapest energy saving techniques.

Second, 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 strongly because their economies tend to depend on climate related production processes like agriculture, fishery and forestry (IPCC (2001)).⁷ The high damage of EU is due to a high potential GDP; the low damage estimate for FSU is due to some expected benefits from moderate temperature increase, like the increased availability of arable land.

Third, 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 . In every 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.

Fourth, 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, implying cost-effectiveness within coalition S .

2.4. Properties of Valuations

To conclude this section, we discuss two general properties which determine the incentive structure in the coalition formation game. Both properties hold for the valuations derived from our simulation model CWSM. 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 .

⁷ Because ROW comprises many heterogenous countries, the climate change damage parameters have been revised downward in all coalition structures in which ROW acts as a singleton following Nordhaus and Yang (1996). Without this modification, non-cooperative emission control rates by ROW would be unrealistically high.

Definition 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*”. This makes 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.

The second property is called positive spillovers, meaning that if country j joins coalition S , all countries that do not belong to $S \cup \{j\}$ are better off.

Definition 2: Positive Spillovers

A coalition game exhibits positive spillovers if and only if for all $S \subset N$, $j \notin S$ and all $l \notin S \cup \{j\}$: $v_l(S \cup \{j\}) > v_l(S)$.

Consequently, there is an “external gain” or a positive spillover from cooperation, making free-riding (i.e. not joining S) 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.⁸

It is then clear that a region’s decision to join a coalition – as well as the stability of an IEA – depends on the magnitude of both properties. On the one hand, the intensity of the superadditivity effect which, together with the sharing rule of the coalitional gain, determines the payoff of cooperation. On the other hand, the intensity of the positive externality effect which determines the free riding payoff. We will study these effects in more detail in

⁸ The importance of positive (and negative) spillovers has been stressed in the recent literature on coalition theory (Bloch 2003, Yi 2003 and Maskin 2003).

section 3, 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 raised because both the coalition members (through superadditivity) and the outsiders (through positive spillovers) benefit.

Table 1 illustrates the magnitudes at stake for the CWSM model. It displays – for a selection of coalitions – global welfare and two environmental variables (carbon concentration and global emissions) in absolute and relative terms. 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 normalization. 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 as suggested by Barrett (1994). In the context of our model with heterogeneous players, the identity of members is more important. 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 are characterized by a relatively flat emission abatement cost curve; hence they can provide cheap abatement. However, apart from emission abatement costs, also environmental damages matter. If a given coalition maximizes its joint welfare, the higher the marginal damages of coalition members are, the higher joint abatement efforts will be (and the higher a positive spillovers), all else being equal. This explains the global importance of EU for cooperation.

Summarizing, cooperation between countries with low marginal abatement costs (i.e. cheap suppliers of abatement effort) and countries with high marginal willingness to avoid climate change (i.e. high demand for abatement effort), proves important to achieve substantial emission reductions in an IEA.⁹

3. Stable Coalitions

In this section, we analyze coalition formation both under the assumption of no transfers and various forms of transfers. Results and arguments are illustrated with the CWSM model. In subsection 3.1, we start with no transfers. Subsection 3.2 deals with ex-ante transfers and shows the advantage of optimal transfers, as defined in Eyckmans and Finus (2004b), compared to simple transfer schemes, as assumed in most papers up to now. Subsection 3.3 analyzes the expansion of coalitions via ex-post transfers.

3.1 No Transfers

In the case of no transfers, there is no coalition that is both internally and externally stable at the same time, see Table 2. In the absence of transfers, although cooperation may be coalitionally and globally 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, regions facing a flat marginal abatement cost and damage cost curve are typically losers from cooperation without transfers. These regions are required to contribute substantial abatement efforts but they do not value strongly the ensuing improvement in environmental quality. Thus, without transfers, China faces the strongest free-rider incentives among all six regions in CWSM.

⁹ Similar conclusions can also be found in Buchner et al. (2002).

3.2 Ex-Ante Transfer Schemes

In the light of the above negative conclusion, we now turn to 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 according to a particular transfer scheme. 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 simple and optimal transfer schemes, the full potential of an optimal design of transfers will become apparent.

3.2.1 Simple Transfer Schemes

All three simple transfer schemes that we consider originate from cooperative coalition theory. It 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 any subcoalition of N (Aumann and Drèze 1974). 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:

$$\hat{v}_i^{SV} = \sum_{\substack{T \subseteq S \\ i \notin T}} \frac{t!(s-t-1)!}{s!} \left[\sum_{k \in T \cup \{i\}} v_k(T \cup \{i\}) - \sum_{k \in T} v_k(T) \right] \quad \forall i \in S \quad (4)$$

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 average marginal contribution over every possible subcoalition T of S (term between square brackets in (4)), weighted by the probability that this subpartition forms (first factor in (4)).

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

$$\hat{v}_i^{NB} = v_i(\{i\}) + \frac{1}{S} \left[\sum_{j \in S} v_j(S) - \sum_{j \in S} v_j(\{i\}) \right] \quad \forall i \in S \quad (5)$$

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 (Chander and Tulkens 1997) in the version as applied in Eyckmans and Tulkens (2003):

$$\hat{v}_i^{CT} = v_i(\{i\}) + \frac{D'_i}{\sum_{j \in S} D'_j} \left[\sum_{j \in S} v_j(S) - \sum_{j \in S} v_j(\{i\}) \right] \quad \forall i \in S \quad (6)$$

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 (and hence the term in square brackets is positive).

The results of our simulations, as summarized in Table 2, 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. First, we cannot generally conclude that simple transfer schemes would always enhance stability. Second, we cannot conclude that the Nash bargaining solution performs better than the other transfer schemes. Though in our example, stable coalition no. 16 generates higher global welfare than in any other stable coalition under the other two transfer schemes, other parameter values or

other models could yield different conclusions. Finally, and most importantly, we have no information about whether we could do any better than the simple transfer schemes considered here and if so what the “best” transfer scheme would be. This question will be addressed in the following section.

3.2.2 Optimal Transfer Schemes

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

Definition 3: 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 PIS, this means that it has sufficient resources to guarantee each of its members at least its free riding payoff. Hence, if S is PIS there does always exist a transfer scheme that can ensure internal stability to all members of S . All what is required to construct such a transfer scheme is to give every member of S at least his free-rider payoff or outside option $v_i(S \setminus \{i\})$. The remaining surplus can then be shared in a proportional way:

Definition 4: Optimal Transfer Schemes (OPTS)

A transfer scheme is called optimal if it satisfies

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

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

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 $\lambda_i(S)$. As long as the surplus of cooperation over the free-riding valuations, as well as all weights $\lambda_i(S)$, are positive, the resulting allocation will be internally stable regardless of the choice of weights. This means that the set of internally stable coalitions does not depend on the choice of weights. Interesting, though less evident, this robustness result also carries over to the set of *externally* stable coalitions. Intuitively, through OPTS a link between internal and external stability is established. If S is PIS, OPTS makes it internally stable and hence all coalitions S without a member i are externally unstable. If S not PIS, S will not be internally stable and S without i is externally stable.

Result 1: Invariance of stability for surplus sharing weights

Any family of weights $\lambda_i(S)$ of an OPTS will lead to the same set of internally and externally stable coalitions after transfers.

A second result is that the coalition with the highest global welfare among all PISC, say coalition S^* , is not only internally stable (because it is PISC) but also externally stable and hence stable.

Result 2: Welfare optimality and stability

Under OPTS, the coalition with the highest global welfare among all PISC is stable.

Intuitively, if S^* were not externally stable, then there would exist a larger PISC with higher global welfare (because of superadditivity and positive spillovers), violating the initial assumption that S^* generates the highest global welfare.

Result 1 and 2 are proved in Eyckmans and Finus (2004b). They are illustrated in Table 2 with our simulations. First, the three simple transfer schemes lead to a different set of stable coalitions. In particular, despite the Nash Bargaining solution and the Chander and Tulkens's transfer scheme only differ in weights, they produce very different results. In contrast, any set of weights for OPTS leads to the same set of stable coalitions as displayed in Table 2. Second, OPTS exploits the full potential of self-enforcing cooperation. In our example, OPTS achieves 94.5 percent of total maximum welfare for coalition no. 5 {USA, EU, CHN, ROW} whereas for the simple transfer schemes the maximum is achieved under the Nash Bargaining solution with only 68.7 percent.

3.3 Ex-Post Transfer Schemes

In this subsection, we consider ex-post transfers. This means that after a coalition S 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 without transfers or with some ex-ante transfer scheme. In our example, these are the coalitions listed in Table 2. 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 3.3.1); in the second case, an outsider $j \notin S$ uses transfers to bribe another outsider $k \notin S$ to join coalition S (subsection 3.3.2).

3.3.1 Expansion of Coalitions through Internal Means

The standard procedure to analyze the expansion of coalitions through internal means is to pick a stable coalition S as starting point and to check whether expansion of this coalition is possible when current coalition members pay some outsider j for joining them (Botteon and Carraro 1997). It is natural to require that expansion of coalition S is feasible if and only if

1) the expansion constitutes a Pareto-improvement to all members of S and for the newcomer j , and 2) the enlarged coalition $S \cup \{j\}$ is internally stable.

The first requirement of a Pareto improvement follows from the presumption that current members of S will only bribe an outsider to join if they are better off after the expansion. Likewise, region j will only accept the deal if it can improve itself. The second requirement simply follows from the definition of stability. Now, it turns out that the first requirement is the condition of superadditivity, which holds in our global emission game (as in many other models), and hence it is non-binding. For the second requirement it can be shown that a necessary condition is that the enlarged coalition $S \cup \{j\}$ is potentially internally stable. Since coalition S^* with the highest global welfare among all PISC is stable under (ex ante) OPTS, this means that no expansion via internal means is possible from this coalition to form $S^* \cup \{j\}$.

Result 3: OPTS exhausts all possibilities for expansion by internal means

Coalition S^ which achieves the highest global welfare among all PISC cannot be improved upon by expansions using internal means.*

Result 3 is proved in the Appendix. It implies that coalition S^* can be regarded as the best possible outcome that can be achieved by means of expansion using internal means. This can also be interpreted that if an optimal ex ante transfer scheme is employed, the analysis of ex post transfers to enlarge a coalition with internal means is redundant. In our example, this means that we cannot do better than coalition no. 5, i.e. {USA, EU, CHN, ROW}.

3.3.2 Expansion of Coalitions 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 obvious, although not exclusively, potential candidates for expansions via external means.¹⁰

Again, two conditions must be satisfied that expansions is successful. First, expansion must constitute a Pareto-improvement to all involved players, which we know is not binding if superadditivity holds. Second, the enlarged coalition $S \cup \{j\}$ must be internally stable. Since by assumption enlargement starts from S which is internally and externally stable, 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,

$$\sum_{i \in S \cup \{j\}} v_i(S \cup \{j\}) + \sum_{i \in S \cup \{j\}} t_i \geq \sum_{i \in S \cup \{j\}} v_i(S \cup \{j\} \setminus \{i\}) \quad (7.a)$$

$$v_k(S \cup \{j\}) + t_k \geq v_k(S) \quad (7.b)$$

must hold where typically $t_i \geq 0$ 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 spillover 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 $S \cup \{j\}$. In general, we cannot derive conclusive theoretical results using general properties like superadditivity or positive spillovers for expansions using external means. Therefore, conditions (7.a) and (7.b) have to be checked for all possible expansions starting from a PIS coalition.

Result 4: Possibility of expansion using external means

Starting at a stable coalition S , expansion using external means of S is possible only if the positive spillover to the external player $k \in N \setminus S$ that pays another outsider $j \in N \setminus S$ to join

¹⁰ As long as expansion via internal means is possible, it is very unlikely that expansion via external means are conducted. This is because outsiders benefit from expansion via internal means through positive spillovers without any contribution, knowing that this expansion is in the interest of all current members of coalition S .

S, is greater than the increase in the free-rider effect of the members of the expanded coalition $S \cup \{j\}$.

Figure 1 shows all possibilities of expansion from coalitions with four members that are stable in our example. Figure 1 illustrates that from coalition no. 5, which is the coalition with the highest global welfare 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. The maximum “degree of optimality” which can be achieved by means of a clever transfer strategy is 95.6 percent (coalition no. 4 comprising all countries except USA). Of course, full participation cannot be achieved since expansion via external means requires that at least one country remains an outsider.

More generally, our results stress that — in the context of incentives to free-ride — 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 other outsiders to become a formal member of an IEA and to comply with its requirements. This result may be particularly useful for the design of future IEAs like for instance post-Kyoto negotiations.

4. Summary and Conclusions

The non-cooperative game theoretic 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 given that an IEA has to be internally and externally stable. 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 shown that ex-post internal transfers are redundant if ex-ante transfers are optimally 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). All results depend only on very general properties of the underlying valuation functions, in particular, only on the properties superadditivity and positive spillovers, which hold for most public good models, not only those that deal with environmental issues.

To show the relevance of the theoretical results, we used a stylized integrated assessment model of climate change policy. Our simulation results highlighted the difficulty of reaching an effective agreement on climate policy because of the large asymmetries across countries and incentives to free-ride on the provision of the public good “climate change control”. We showed that the situation can be largely improved by adopting an appropriate transfer scheme. In the case of transfer using internal resources, our “optimal” transfer scheme leads to far better results than the “simple” transfer schemes well-known in existing literature. In the case of transfers using external resources, we demonstrated that it is possible to reach a self-enforcing agreement involving all but one country but where this outsider contributes financially, in its self interest, to stabilize a successful agreement.

Both results have important implications for post-Kyoto climate policy negotiations. On the one hand, they stress how important is to devote sufficient time to the design of transfer schemes so that participation can be raised in the future. On the other hand, they highlight the important role that current outsiders like the USA and Australia 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 promote energy efficiency and, eventually perhaps to assume quantitative emission ceilings and to become a full member of the Kyoto Protocol. These transfers may well be in the interest of the outsiders 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.

The analysis of this paper could be extended in several directions. For instance, it would be interesting to analyze the role of transfers when multiple coalitions can form or when coalitions instead of single players could deviate from an existing coalition. Moreover, the relationship between the optimal transfer scheme proposed in this paper and the transfer scheme implicit in the implementation of an emission trading market is certainly also a research topic for future research (see Bosello et al. 2004 for some preliminary numerical results).

References

- Altamirano-Cabrera, J.-C. and Finus, M. (2006). Permit trading and stability of international climate agreements. *Journal of Applied Economics*, 9 (1): 19-48
- Aumann, R.J. and Drèze J.H. (1974). Cooperative games with coalition structures. *International Journal of Game Theory* 3 (4): 217–237.
- Barrett, S. (1994). Self-enforcing international environmental agreements. *Oxford Economic Papers* 46 (October): 804–878.
- Bloch, F. (2003). Non-cooperative models of coalition formation in games with spillovers. In C. Carraro (Ed), *Endogenous Formation of Economic Coalitions*, 35–79. Cheltenham, UK: Edward Elgar.
- Bosello, F., Buchner, B. and Carraro, C. (2003). Equity, development, and climate change control. *Journal of the European Economic Association* 1 (2-3): 601–611.
- Bosello, F., Buchner, B., Carraro, C. and Raggi, D. (2004). Can equity enhance efficiency? Some lessons from climate negotiations. In C. Carraro and V. Fragnelli (Eds.), *Game practice and the environment*, 37-64. Cheltenham, UK: Edward Elgar.
- Botteon, M. and Carraro, C. (1997). Burden-sharing and coalition stability in environmental negotiations with asymmetric countries. In C. Carraro (Ed), *International environmental negotiations: strategic policy issues*, 26–55. Cheltenham, UK: Edward Elgar.
- Buchner, B, Carraro, C. and Cersosimo, I. (2002). Economic consequences of the US withdrawal from the Kyoto Protocol. *Climate Policy* 76 (2): 1–20.
- Carraro, C. (2000). Roads towards international environmental agreements. In H. Siebert (Ed), *The economics of international environmental problems*, 169–202. Tübingen: Mohr Siebeck.
- Carraro, C. and Siniscalco, D. (1993). Strategies for the international protection of the environment. *Journal of Public Economics* 52 (3): 309–328.
- Carraro, C. and Siniscalco, D. (1998). International environmental agreements: incentives and political economy. *European Economic Review* 42 (3-5): 561–572.
- Carraro, C. and Siniscalco, D. (2001). Transfers, commitments and issue linkage in international environmental negotiations. In A. Ulph (Ed), *Environmental policy, international agreements and international trade*, 19-37. Oxford: Oxford University Press.
- Chander, P. and Tulkens, H. (1997). The core of an economy with multilateral environmental externalities. *International Journal of Game Theory* 26 (3): 379–401.
- d’Aspremont, C., Jacquemin, A., Gabszewicz J. J. and Weymark, J. A. (1983). On the stability of collusive price leadership. *Canadian Journal of Economics* 16 (1): 17–25.
- Eyckmans, J. and Finus, M. (2004a). Measures to enhance the success of global climate treaties. Forthcoming in *International Environmental Agreements: Politics, Law and Economics*.

- Eyckmans, J. and Finus, M. (2004b). An almost ideal sharing scheme for coalition games with externalities. CLIMNEG Working Paper Nr. 62, Leuven, Belgium: Katholieke Universiteit Leuven.
- Eyckmans, J., and Finus, M. (2006). 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.
- Eyckmans, J. and Tulkens, H. (2003). Simulating coalitionally stable burden sharing agreements for the climate change problem. *Resource and Energy Economics* 25 (4): 299–327.
- Finus, M. (2001). *Game theory and international environmental cooperation*. Cheltenham, UK: Edward Elgar.
- Finus, M. (2003a). Stability and design of international environmental agreements: the case of transboundary pollution. In H. Folmer and T. Tietenberg (Eds.), *International yearbook of environmental and resource economics*, 2003/4, 82–158. Cheltenham, UK: Edward Elgar.
- Finus, M. (2003b). New developments in coalition theory: an application to the case of global pollution. In L. Marsiliani, M. Rauscher and C. Withagen (Eds.), *Environmental policy in an international perspective*, 19–49. Dordrecht, The Netherlands: Kluwer.
- Finus, M., Sáiz, E. and Hendrix, E.M.T. (2004). An empirical test of new developments in coalition theory for the design of international environmental agreements. *Mansholt Working Paper Series*, Nr. 14, Wageningen, The Netherlands: University of Wageningen.
- Hoel, M. (1992). International environment conventions: the case of uniform reductions of emissions. *Environmental and Resource Economics* 2 (2): 141–159.
- IPCC (2001). *Climate change 2001: mitigation. Contribution of working group III to the third assessment report of the intergovernmental panel on climate change*, Cambridge, UK: Cambridge University Press.
- Jeppesen, T. and Andersen, P. (1998). Commitment and fairness in environmental games. In N. Hanley and H. Folmer (Eds.), *Game theory and the environment*, 65–83. Cheltenham, UK: Edward Elgar.
- Maskin, E. (2003). Bargaining, coalitions, and externalities. Mimeo. Paper presented at the EEA-ESEM Meeting in Stockholm, August.
- Nordhaus, W.D. and Yang, Z. (1996). A regional dynamic general-equilibrium model of alternative climate-change strategies. *American Economic Review* 86 (September): 741–765.
- Petrakis, E. and Xepapadeas, A. (1996). Environmental consciousness and moral hazard in international agreements to protect the environment. *Journal of Public Economics* 60 (1): 95–110.
- Weikart, H.-P., Finus, M. and Altamirano-Cabrera, J.-C. (2006). The impact of surplus sharing on the stability of international climate coalitions. *Oxford Economic Papers*, 58 (2): 209–232

Yi, S.-S. (2003). Endogenous formation of economic coalitions: a survey of the partition function approach. In C. Carraro (Ed), Endogenous formation of economic coalitions, 80–127. Cheltenham, UK: Edward Elgar.

Appendix

The first requirement of Pareto-improvement means that

$$\forall i \in S: \quad v_i(S \cup \{j\}) + t_i \geq v_i(S) \quad (\text{A1.a})$$

$$v_j(S \cup \{j\}) + t_j \geq v_j(S) \quad (\text{A1.b})$$

must hold where typically $t_i \leq 0$ and $t_j \geq 0$. By adding (A1.a) and (A1.b), summing over all regions $S \cup \{j\}$ involved in the expansion and noting that transfers should balance (i.e.

$\sum_{i \in S \cup \{j\}} t_i = 0$), we obtain: $\sum_{i \in S \cup \{j\}} v_i(S \cup \{j\}) \geq \sum_{i \in S \cup \{j\}} v_i(S)$ which implies that superadditivity

should hold in the move from coalition S to $S \cup \{j\}$. The second requirement of internal stability of the expanded coalition means:

$$\forall i \in S: \quad v_i(S \cup \{j\}) + t_i \geq v_i(S \cup \{j\} \setminus \{i\}) \quad (\text{A2.a})$$

$$v_j(S \cup \{j\}) + t_j \geq v_j(S) \quad (\text{A2.b})$$

A similar manipulation as above reveals that a necessary condition (2.a) and (2.b) to hold is

$\sum_{i \in S \cup \{j\}} v_i(S \cup \{j\}) \geq \sum_{i \in S \cup \{j\}} v_i(S \cup \{j\} \setminus \{i\})$ which is the condition of potential internal stability

(PIS).

Table 1: Welfare and Environmental 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
...	...	...	...	...	...
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: $\left(\sum_{i=1}^N (w_i(c^p) - w_i(c^N))\right) / \left(\sum_{i=1}^N (w_i(c^F) - w_i(c^N))\right)$ 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\$₁₉₉₀ (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\$₁₉₉₀ 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*

	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
OPTS	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").