

CTN (Venice, January 2008)

Sticks and carrots for the design of international climate agreements with renegotiations

Hans-Peter Weikard* and Rob Dellink

Environmental Economics and Natural Resources Group, Wageningen University

Abstract: This paper examines stability of international climate agreements for carbon abatement under an optimal transfer rule and renegotiations. The optimal transfer rule suggested to stabilise international environmental agreements (Eyckmans and Finus 2004, Weikard 2005) is no longer optimal when agreements are renegotiated. We determine the necessary and sufficient conditions for optimal self-enforcing sequences of agreements. If these conditions are met, then transfer payments can be arranged such that no country wants to change its membership status at any stage. In order to demonstrate the applicability of our condition we use the STACO model, a 12-regions global model, to assess the impact of well-designed transfer rules on the stability of an international climate agreement. Although there are strong free-rider incentives, we find a stable grand coalition in the first commitment period in a game with one round of renegotiations.

JEL classification: C72, D62, H41, H77

Keywords: Stability of coalitions, international environmental agreements, partition function approach, sharing rules, optimal transfers

* Presenter: Department of Social Sciences, Wageningen University
Hollandseweg 1, NL-6706 KN Wageningen, The Netherlands.
Email: Hans-Peter.Weikard@wur.nl

1 Introduction

Efficient management of the global commons such as stability of climatic conditions and the conservation of biodiversity requires a grand coalition of the nation states responsible for environmental regulation and legislation. However, the grand coalition will generally not be stable as individual countries have an incentive to take a free-rider position in the provision of such global environmental goods. This paper examines the options to establish a stable international climate agreement (ICA) that comes as close as possible to the first best – but generally unstable – grand coalition. The notion of stability we employ is crucial. We use the concept of internal and external stability. A coalition is said to be stable if and only if no coalition member has an incentive to leave (internal stability) and no non-member has an incentive to join (external stability); *cf.* d'Aspremont *et al.* (1983). This notion of stability is applicable when the membership decision is taken only once (and for all). As we will relax this assumption and allow for renegotiations, we will introduce an appropriate refinement of the stability concept. This is discussed in section 2 below.

There exists by now an extensive literature on the stability of international environmental agreements. Here we focus on games where a single agreement is proposed that can be signed or not. This class of games is usually called cartel games with open membership. Different solution concepts have been employed to analyse this type of game. Hoel (1992) considers a 'constrained social optimum' where social net benefits from greenhouse gas (GHG) abatement are maximised subject to an individual rationality constraint that requires that coalition members gain compared to the non-cooperative equilibrium. Eyckmans (1999) examines coalition stability for uniform abatement efforts using a similar individual rationality constraint. Individual rationality is also a condition of the core solutions considered by e.g. Chander and Tulkens (1995). It is implicitly assumed that a coalition would completely dissolve if a coalition member leaves. By contrast the notion of (internal and external) stability implicitly assumes that the remaining coalition stays intact if a coalition member leaves. Only single deviations are considered. One might call this the pessimistic view where deviations occur whenever they are immediately profitable. Solution concepts relying on individual rationality are more optimistic in the sense that players are assumed to always stick to a coalition if their coalition payoff exceeds their non-cooperative payoff. In this paper we take a pessimistic – or cautious – view.

We explore the stability of international climate agreements. Each player (countries or regions) chooses whether or not to sign the agreement and become a member of a unique coalition (the climate agreement). Then the coalition and the remaining singletons fix their abatement levels playing a Nash-Cournot game. As abatement is a pure public good the equilibrium of the abatement game is generally inefficient. Only the coalition of all, the grand coalition, would overcome the inefficiency. The literature on cartel membership games has emphasised, however, that large coalitions, including the grand coalition, are only stable if

gains from cooperation are small, that is, whenever coalition formation is rather unimportant. If gains from cooperation are large, the analysis of cartel games shows that only small coalitions consisting of no more than 2 or 3 players are stable (e.g. Barrett 1994, Finus and Rundshagen 2003). These results have been obtained for models with identical players and do not generalise when players differ (see Weikard 2005). Our results in this paper are also in contrast to these earlier findings. We show that there are options to stabilise successful ICAs. There are three main drivers of our result. First, we consider asymmetric countries. With asymmetric countries, in particular when countries differ in abatement cost, a coalition can exploit cheap abatement options if a low cost country joins the coalition. Second, we consider transfers between coalition members. Transfers set incentives for low cost countries to join the coalition. We employ a transfer rule suggested by Eyckmans and Finus (2004), Weikard (2005) and McGinty (2007) that guarantees internal stability whenever it is feasible given the payoffs. Third, we allow for renegotiations of the agreement. Renegotiations improve the incentives to join a coalition. Players may be forced to cooperate at the first stage if there is a credible threat to be punished at the second stage. Moreover the second stage payoffs can be used to reward members of the first stage agreement. Similar to our approach Rubio and Ulph (2007) study on renegotiations of international environmental agreements and dynamic membership. However, their approach is confined to symmetric countries and does not capture abatement cost differences and the impact of transfers to stabilise agreements.

An important determinant of the stability of a coalition is the sharing of the coalition payoff among members. In the debate on climate change policies a number of different sharing rules for international climate agreements have been suggested; see Rose et al. (1998). Only few studies have addressed the impacts of different sharing rules on the stability of international climate agreements. Bosello et al. (2003) consider sharing rules for abatement costs, Altamirano-Cabrera and Finus (2006) consider sharing rules for tradable emission permits and Weikard, Finus and Altamirano-Cabrera (2006) consider sharing rules for the gains from cooperation. The differences are crucial. If sharing rules are applied to costs or permits, there is no guarantee that payoffs satisfy the individual rationality constraint. By contrast, if sharing is applied to the gains from cooperation, individual rationality is always satisfied as long as a coalition is at all profitable, i.e. it achieves at least the sum of what its members can achieve without cooperation. Eyckmans and Finus (2004), Weikard (2005) and McGinty (2007) go one step further. They propose to apply a sharing rule to the difference of the coalition payoff and the sum of the outside option payoffs of coalition members. The outside option payoff is the payoff a player would receive when leaving the coalition. It is obvious that the internal stability condition can only be met when the coalition payoff exceeds the sum of the outside option payoffs. Here we extend this reasoning to a multiperiod setting where the agreement can be renegotiated in each period.

To study the effects of optimally designed sharing rules in an empirical setting and to illustrate their impacts we examine the stability of ICAs. We do this using the STACO model

introduced by Finus et al. (2006) and refined by Nagashima et al. (2006). STACO is a global model comprising 12-regions for each of which abatement cost and benefit functions are defined. The empirical analysis in this paper extends the work of Weikard, Finus and Altamirano-Cabrera (2006) who consider results from the STACO model and compare proportional sharing of gains from cooperation according to past emissions, regional income, population and other indicators. They find that an emission based proportional sharing rule performs best in the sense that it stabilises a coalition that gives higher global net benefits and lower stocks of carbon pollution than any other of the examined rules. However, as only a set of eight rules was examined in that study, it remained an open question whether coalitions can be stabilised that perform even better. Our analysis in this paper shows that an optimal transfer rule gives significantly better results. If we do not consider renegotiations, we find that there exist stable coalitions of up to 6 out of 12 regions. The best performing stable coalition comprises 5 regions and achieves about 46% of the gains the Grand Coalition could achieve as compared to non-cooperation. With renegotiations, considering two commitment periods, we even find a stable grand coalition in the first commitment period of 20 or 30 years, followed by a 5 regions coalition for the second commitment period. These agreements achieve 59% of the potential gains over a century.

The paper is organised as follows. The next section introduces optimal sharing rules for a multi-period setting. Section 3 provides a brief overview of empirical specifications of the STACO model for our case of ICAs. Section 4 presents the results. Section 5 concludes and discusses implications for climate policies.

2 Optimal sharing in renegotiation proof coalitions

We consider an open membership cartel formation game with renegotiations. That is, we consider several but finitely many periods (commitment periods) in each of which a single coalition or cartel (the ICA) can form. In each period $t = 1, \dots, T$ individual countries $i \in N$ decide to join or not to join the agreement. Payoffs that accrue in period t depend on the coalition formed in that period $K^t \subseteq N$. Payoffs are determined from costs and benefits of equilibrium abatement of GHGs in an abatement game that is specified in the next section. Here it is sufficient to note that GHG abatement is a public good. It is well known for public goods games that the grand coalition $K = N$ is efficient, while any partial agreement with $K \subset N$ is not (e.g. Dasgupta 1982). In a public goods game a singleton coalition $K = \{i\}$ will not be effective and give the same payoffs as $K = \emptyset$. We will refer to both cases as ‘all-singletons’. To be more precise about payoffs, the abatement game in period t invokes a partition function $V^t(K^t)$ that determines payoffs V_K^t for the coalition and for each singleton player $V_j^t(K^t)$, $j \notin K^t$ in period t . Before we can introduce our notion of stability (of a sequence of coalitions) we need to define individual payoffs for coalition members. We

assume that some sharing rule r applies that distributes the coalition payoff $V_K(K)$ among members. Thus we arrive at a per-member partition function, also called a valuation function. For convenience we denote it by $V({}^r K)$ and individual payoffs under coalition K when sharing rule r applies are denoted by $V_i({}^r K)$. Of course, for every sharing rule r we have $V_K(K) = \sum_{i \in K} V_i({}^r K)$. We adopt the shorthand notation K_{-i} for $K \setminus \{i\}$ with $i \in K$ and K_{+j} for $K \cup \{j\}$ with $j \notin K$. Define the coalition surplus $S_K \equiv V_K(K) - \sum_{i \in K} V_i(K_{-i})$. We call $V_i(K_{-i})$ i 's outside option payoff or defection payoff. Note that it is, of course, independent of the sharing rule applied by coalition K_{-i} .

Now consider period T and consider the following class of sharing rules for coalition K^T :

Each coalition member $i \in K^T$ receives

$$V_i^T({}^r K^T) = V_i^T(K_{-i}^T) + s_i \text{ for all } i \in K^T, \quad (1)$$

$$\text{with } \sum_{i \in K^T} s_i = S_{K^T} \text{ and}$$

$$\text{for all } i \in K^T, s_i \geq 0 \text{ if and only if } S_{K^T} \geq 0.$$

By construction, sharing according to (1) – optimal sharing – gives each coalition member at least its outside option payoff if this is feasible, i.e. $S_{K^T} \geq 0$. Hence, optimal sharing gives priority to setting incentives to join the agreement. Weikard (2005) shows that every coalition that is internally stable under some arbitrary sharing rule will also be internally stable under the family of sharing rules characterised by (1). The underlying standard notion of stability is as follows:

DEFINITION 1 A coalition K is stable under sharing rule r , if it is

(i) internally stable, i.e.

$$V_i^r(K) \geq V_i(K_{-i}) \text{ for all } i \in K,$$

and

(ii) externally stable, i.e.

$$V_j(K) > V_j^r(K \cup \{j\}) \text{ for all } j \in N \setminus K.^1$$

In a setting with renegotiations, the notion of stability of Definition 1 has to be refined. First, strictly speaking, the object of stability is no longer a coalition under a sharing rule $r, {}^r K$ but rather sequence of coalitions coupled with sharing rules $\langle {}^{r^1} K^1, \dots, {}^{r^T} K^T \rangle$. We adopt the

¹ The tie-breaking rule is here that a player would join the coalition if she is indifferent between joining and staying out. Hence, by this definition, the empty set \emptyset is not externally stable, and a trivial coalition is internally stable.

terminology that coalition K^t is *generalised stable* if no member wants to leave and no non-member wants to join. It is important to note that generalised stability of a coalition in a given period may depend on the coalitions formed in other periods and Definition 1 does not apply. Furthermore, renegotiations rule out the play of a Pareto dominated subgames. The equilibrium concept employed is *renegotiation proof equilibrium* introduced by Bernheim et al. (1987), Bernheim and Ray (1989), and Farrell and Maskin (1989).

Before formally defining renegotiation proof sequences of coalitions, we need to introduce some further notation and a definition of generalised stability.

Denote by ${}^w\tilde{K}_i^T$ the stable coalition with sharing rule w that gives the worst payoff of all stable coalitions to player i in period T , that is $V_i^T({}^w\tilde{K}_i^T) \leq V_i^T({}^rK^T)$ for all stable ${}^rK^T$.

We can now recursively define generalised stability. A definition of is renegotiation proof equilibrium follows.

DEFINITION 2 (i) In the last period a coalition K^T is generalised stable under sharing rule ${}^r{}^T$ if and only if it is stable.

(ii) A sequence of coalitions $\langle {}^{r^{T-1}}K^{T-1}, {}^{r^T}K^T \rangle$ is generalised stable if and only if

${}^{r^T}K^T$ is generalised stable,

and for all $i \in K^{T-1}$

$$V_i^{T-1}({}^{r^{T-1}}K^{T-1}) + V_i^T({}^{r^T}K^T) \geq V_i^T(K_{-i}^{T-1}) + V_i^T({}^w\tilde{K}_i^T), \quad (2)$$

and for all $j \in N \setminus K^{T-1}$

$$V_j^{T-1}(K^{T-1}) + V_j^T({}^{r^T}K^T) > V_j^{T-1}({}^{s^{T-1}}K_{+j}^{T-1}) + V_j^T({}^w\tilde{K}_j^T). \quad (3)$$

(iii) For longer sequences conditions for generalised stability can be established by working backwards from (2) and (3). Let W_i^t be the worst generally stable coalition sequence for player i .

A sequence of coalitions of length $T-t+1$, i.e. $\langle {}^{r^t}K^t, \dots, {}^{r^T}K^T \rangle$, is generalised stable

if the subsequence of length $T-t$, i.e. $\langle {}^{r^{t+1}}K^{t+1}, \dots, {}^{r^T}K^T \rangle$, is generalised stable and

if for all $i \in K^t$

$$V_i^t({}^{r^t}K^t) + \sum_{\tau=t+1}^T V_i^\tau({}^{r^\tau}K^\tau) \geq V_i^t(K_{-i}^t) + W_i^{t+1},$$

and for all $j \in N \setminus K^{T-1}$

$$V_j^t(K^{T-1}) + \sum_{\tau=t+1}^T V_j^\tau(r^\tau K^\tau) > V_j^t(s^t K_{+j}^t) + W_j^{t+1}.$$

DEFINITION 3 (i) In the last period, coalition K^T is renegotiation proof under sharing rule r^T if and only if it is generalised stable and Pareto undominated by any other generalised stable coalition.

(ii) A sequence of coalitions $\langle r^t K^t, \dots, r^T K^T \rangle$ is renegotiation proof if and only if

it is generalised stable and Pareto undominated by any other generalised stable sequence stretching from t to T

and all its subsequences $\langle r^{t-1} K^{t-1}, \dots, r^T K^T \rangle$, $\langle r^{t-2} K^{t-2}, \dots, r^T K^T \rangle$, ..., $\langle r^T K^T \rangle$ are generalised stable and Pareto undominated.

In the remainder of the paper we confine the analysis to a two-period game, $T = 2$. To extend the analysis to more periods is straightforward, based on Definitions 2 and 3, but tedious.

For $T = 2$, $V_i^2({}^w \tilde{K}_i^2)$ is the worst credible punishment that can be imposed on i in period 2.

Condition (2) requires of a renegotiation proof equilibrium $\langle r^1 K^1, r^2 K^2 \rangle$ that the payoff of a period 1 coalition member (left hand side) is at least as large as the payoff from defecting from coalition K^1 (right hand side). It is a credible threat that ${}^w \tilde{K}_i^2$ will be played as it is a renegotiation proof subgame. Condition (3) requires that a singleton player in period 1 receives at least as much as when entering coalition K^1 . Conditions (2) and (3) are generalisations of the notions of internal and external stability, respectively.

Using definition 2 we are now going to establish a condition that can be used to determine whether for a given partition function we can find valuations (or sharing rules) such that a sequence of coalitions is renegotiation proof. For simplicity we restrict the exposition to a time invariant partition function. Again, generalisations are straightforward but involve more tedious notation. Also we focus on internal stability, i.e. condition (2). The following proposition establishes that external stability, condition (3), is of little concern.

PROPOSITION 1: If K^1 is generalised internally stable and generalised externally unstable such that there exists $j \notin K^1$ with $V_j(K^1) + V_j(r^2 K^2) \leq V_j(r^1 K_{+j}^1) + V_j({}^w \tilde{K}_j^2)$, meaning that j prefers to enter coalition K^1 , then the enlarged coalition K_{+j}^1 will be internally stable.

Proof: The proof is similar to Weikard (2005), Lemma 1. ■

Hence, in what follows we focus on internal stability. From condition (2) we can see that for a coalition K^1 to be stable each coalition member must at least receive her defection payoff

$$D_i(K^1) \equiv V_i(K_{-i}^1) + V_i({}^w \tilde{K}_i^2).$$

Therefore, a necessary condition for stability of K^1 is that the sum of payoffs of the coalition members is at least as large as the sum of their defection payoffs:

$$\sum_{i \in K^1} (V_i({}^{r^1} K^1) + V_i({}^{r^2} K^2)) \geq \sum_{i \in K^1} (V_i(K_{-i}^1) + V_i({}^w \tilde{K}_i^2)). \quad (4)$$

The stability of K^1 depends upon defection payoffs of the members of K^1 and the sharing rule r^2 applied in period 2; the latter only if for some $i \in K^1$ also $i \in K^2$. The defection payoff is fully specified by the partition function and independent of sharing. The partition function determines the payoffs of the singleton players and, hence, the outside option payoff of $i \in K^1$, $V_i(K_{-i}^1)$ is independent of sharing. If player $i \in K^1$ defects, then her period-2 payoff is $V_i({}^w \tilde{K}_i^2)$. If the player is singleton, i.e. $i \notin \tilde{K}_i^2$, the payoff is independent of the sharing rule w . If $i \in \tilde{K}_i^2$, then she receives the worst possible payoff compatible with renegotiation proofness of \tilde{K}_i^2 .

The next step in our argument is to determine how sharing in coalitions K^1 and K^2 can be arranged such that condition (4) is met in all cases where this is possible for a given partition function. To achieve this the sharing rule r^2 applied to K^2 should distribute the coalition surplus S_{K^2} entirely to members of $K^1 \cap K^2$, leaving members of K^2 that are not members of K^1 with their outside option payoff. Intuitively this rewards coalition members of the first period and lowers their incentive for defection. Hence, by construction we have the following:

PROPOSITION 2: (i) Consider a renegotiation proof period-2 coalition K^2 . Suppose sharing in period 2 is arranged such that $\sum_{i \in K^1 \cap K^2} V_i({}^{r^2} K^2) = S_{K^2}$. Then if for such K^2 and r^2 condition (4) is not met, then $\langle {}^{r^1} K^1, {}^{r^2} K^2 \rangle$ cannot be a renegotiation proof equilibrium for any sequence of sharing rules r^1, r^2 .

(ii) Again, let K^2 be renegotiation proof. Suppose sharing in period 2 is arranged as stated under (i). Then if for such K^2 and r^2 condition (4) is met for K^1 , then there exists a sharing rule r^{1*} such that $\langle {}^{r^{1*}} K^1, {}^{r^2} K^2 \rangle$ is a renegotiation proof equilibrium.

Proof: We do not work out a formal proof. Part (i) follows from the construction of s . Part (ii) follows from the fact that in (4) the period-1 coalition payoff $\sum_{i \in K^1} V_i^r(K^1)$ is independent of r

and from the fact that there is no restriction on r . Hence, as long as the (4) holds, the period-1 coalition payoff V_{K^1} can always be distributed in a way to meet the internal stability condition

(2) for each and every $i \in K^1$. ■

Proposition 2 allows to identify the largest (or most successful) renegotiation proof (or stable) coalitions in a coalition formation game for the provision public goods.

3 Empirical model and data

This section and the next provide an application. We identify the best renegotiation proof equilibrium in a greenhouse gas abatement game, that is we identify a sequence of international climate agreements that is stable and that achieves a higher global payoff than other stable agreements. The stage game we consider is standard in this domain (e.g. Carraro and Siniscalco 1993, Barrett 1994, Botteon and Carraro 1997, Rubio and Ulph 2007). It consists of an announcement game followed by an abatement game. We explore this game as a one-shot game (i.e. with one commitment period) and with one round of renegotiations (i.e. with two commitment periods). Formally the stage game is characterised by a set of regions $N = \{1, 2, \dots, n\}$. At the beginning of commitment period t ($t = 1, 2$) each player i chooses from a strategy set $\sigma_i^t \in \{0, 1\}$; $\sigma_i^t = 0$ means that i is not joining the ICA in period t ; $\sigma_i^t = 1$ means that i joins the ICA in period t . As before, we denote by $K^t \subseteq N$ the set of coalition members (signatories) in commitment period t . Now consider a given coalition K^t . Chander and Tulkens (1995) call this a ‘partial agreement’. In the period- t abatement game members of K^t act like a single player and maximise their joint payoff. They adopt GHG abatement paths which are jointly optimal responses to others’ abatement paths. Other players, the singletons, maximise individual payoffs. They adopt abatement paths which are optimal responses to others’ emissions. Generally we denote i ’s GHG abatement path in period t by \bar{q}_i^t . The period- t Nash equilibrium abatement strategy is denoted by $\bar{q}_i^{*t}(K^t)$. The abatement game described generates a partition function that gives payoffs for the coalition and the singletons for every coalition that may form.² Generally, each player receives benefits b_i from total abatement $\bar{q}^t = \sum_{i \in N} \bar{q}_i^t$ and incurs costs c_i for own abatement \bar{q}_i^t . A singleton player under coalition K^t receives

$$V_i(K^t) = b_i(\bar{q}^{*t}(K^t)) - c_i(\bar{q}_i^{*t}(K^t)), \text{ for all } i \in N \setminus K^t. \quad (5)$$

The coalition payoff is given by

$$V_{K^t}(K^t) = \sum_{i \in K^t} (b_i(\bar{q}^{*t}(K^t)) - c_i(\bar{q}_i^{*t}(K^t))). \quad (6)$$

The partition function on which the stability analysis is based arises from the abatement game characterised by (5) and (6).

To further specify the model, we adopt the STACO-model introduced by Finus et al. (2005) and refined by Nagashima et al. (2005) and Weikard, Dellink and van Ierland (2006). Here, we focus on the main features of the model. We consider twelve world regions; USA (USA), Japan (JPN), European Union-15 (EU15), other OECD countries (OOE), Eastern European

² One qualification is needed. The partial agreement abatement game generates a partition function only if it has unique solution. This condition is satisfied in the empirical setting of the STACO model.

countries (EET), former Soviet Union (FSU), energy exporting countries (EEX), China (CHN), India (IND), dynamic Asian economies (DAE), Brazil (BRA) and rest of the world (ROW). We account for benefits from abatement to infinity, but adopt a shorter planning horizon of 100 years, ranging from 2011 to 2110, for determining abatement paths. In this setting the intertemporal aspects of climate change are well reflected. Benefits from abatement are avoided damages which, in turn, depend on stock of CO₂; each region receives a share of global benefits. As described in Dellink et al. (2004), marginal benefits can be assumed to be constant. Marginal abatement costs are assumed to be a quadratic function of the region's own abatement level. Appendix A provides the specification of regional payoffs. The assumption of linear benefits implies that regions have dominant strategies; the optimal response functions do neither depend on the actions of other players, nor on the stock of greenhouse gases. We calibrate the model as far as possible (regional emission and GDP paths, regional abatement costs) on the MIT-EPPA model (Babiker et al., 2001 and Ellerman and Decaux, 1998), but as this model does not contain a damage cost module, we have to rely on other sources (mainly Nordhaus and Boyer, 2001 and Tol, 1997) for calibrating the benefits from abatement; see Dellink et al. (2004) and Nagashima et al. (2005) for more details.

4 Results and discussion

This section presents the results of coalition formation for the STACO game with one commitment period (4.1) and two commitment periods (4.2).

4.1 *Optimal sharing in the one-shot game*

In the one-shot game, coalition members commit to abatement and transfer paths for the full planning period of 100 years. We find 108 stable and pareto undominated, i.e. renegotiation proof, coalitions. In Table 1 we present net present values (NPVs) and resulting CO₂ concentrations for the benchmark cases (All-Singletons and Grand Coalition) and for the best performing stable coalitions.

In our setting with asymmetric players global payoff does not just depend on the number of players in the coalition but also on their characteristics. For instance, coalitions where China is member will generally implement higher abatement levels than coalitions where China is not involved. As China's marginal abatement costs are very low, the efficient division of abatement efforts will involve a substantial contribution of China. In fact, China appears as member in all ten best-performing coalitions. Of course, China's involvement in a coalition requires sufficient availability of transfers – else China would prefer to take a free-rider position. We find China in 50 out of 108 stable coalitions.

Table 2 provides more detailed information for the All-Singletons case (which is stable but dominated) and for the best performing stable coalition. We report abatement levels and payoffs for individual regions. As a reference, abatement in 2011 and discounted payoffs are reported. The best performing coalition consists of the USA, Central and Eastern European Countries in Transition (EET), China, India and the Dynamic Asian Economies (DAE). All regions are better off under partial cooperation than in the non-cooperative situation, showing that the gains from cooperation are substantial.

Table 1. Best performing stable coalitions – one-shot game

Coalition members	NPV of payoff [Billion US\$]	Concentrations 2110 [Gton]
All-Singletons*	5238	1448.3
Grand Coalition*	15211	1304
USA, EET, CHN, IND, DAE	9830	1403.0
EU15, EET, EEX, CHN, IND	9810	1403.2
EU15, OOE, EET, CHN, IND	9701	1404.2
EU15, EET, CHN, IND, DAE	9697	1404.4
USA, EET, EEX, CHN, DAE, BRA	9613	1405.2
EU15, OOE, EET, EEX, CHN, BRA	9486	1406.4
USA, CHN, IND, DAE, BRA	9484	1406.2
EU15, EET, EEX, CHN, DAE, BRA	9484	1406.5
USA, OOE, EET, EEX, CHN	9469	1406.8
EU15, EEX, CHN, IND, BRA	9455	1406.4

* All-Singletons and Grand Coalition are not stable.

Table 2. Key results for selected coalitions – one-shot game

	All-Singletons		Coalition of USA, EET, CHN, IND, DAE			
	Abatement in 2011 [% of BaU emissions]	NPV of payoff [Billion US\$]	Abatement in 2011 [% of BaU emissions]	NPV of payoff [Billion US\$]	NPV of transfer [Billion US\$]	Incentive to change signal (NPV) [Billion US\$]
USA	9.9	1117	13.2	1603	-663	-9
JPN	2.5	943	2.5	1933	0	-207
EU15	7.6	1240	7.6	2595	0	-274
OOE	5.6	188	5.6	386	0	-19
EET	4.4	71	28.7	138	64	-1
FSU	6.7	362	6.7	749	0	-59
EEX	1.9	164	1.9	336	0	-15
CHN	14.8	298	50.1	415	452	-2
IND	10.5	268	37.8	485	83	-3
DAE	1.9	136	18.1	261	64	-1
BRA	0.1	84	0.1	172	0	-4
ROW	6.3	365	6.3	755	0	-60
Global	8.0	5238	16.9	9830	0	-654

The member with the highest marginal benefits from abatement in the best performing equilibrium is the USA; thus, the transfer scheme involves a payment from the USA to the other regions, as shown in Table 2. The low marginal abatement costs in China induce a huge abatement effort in China and the associated costs are largely funded by the transfer scheme.

The main beneficiaries from this coalition are the two outsiders: Japan and the European Union benefit much from the coalition's increased abatement, compared to the All-Singletons case. They do not increase their abatement efforts, however. This reflects the public good characteristic of abatement. Thus, their incentive to join the coalition (to change membership) is large and negative.

4.2 Optimal sharing in the renegotiation game

Renegotiations after 50 years

As discussed in the introduction, renegotiations may influence the outcome as players can change their membership decision. Introducing a round of renegotiations after 50 years may induce larger coalitions in the first stage, as non-cooperative behaviour may be punished by playing a worse equilibrium in the second stage.

As Table 3 shows, there are multiple equilibria in the second stage. This opens possibilities for punishment and brings about a huge number of equilibria in the first stage. We find a total of 35.671 renegotiation proof equilibria. Table 3 reports NPVs of payoffs for the 10 best performing renegotiation proof coalitions (RPC). The table offers a number of interesting observations.

First, although the second stage of the game has a shorter time horizon than the one-shot game, the best performing equilibria are largely the same. The best performing equilibria in the first stage are, however, more ambitious. Where in the one-shot game 46% of the potential gain from cooperation may be reaped, the two-stage game can close the gap between All-Singletons and the Grand Coalition by 57%. Clearly, the combination of renegotiations and a scheme of optimal transfers is quite powerful. In comparison, a two-stage game with an exogenous transfer rule can only reap 25% of the potential gain (*cf.* Weikard, Dellink and van Ierland 2006).

Secondly, one of the most ambitious RPCs comprises USA, EU15, EET, EEX, China and India in the first stage and the same coalition without the USA in the second stage. As the USA have strong incentives not to be involved in the second stage coalition, it can be persuaded to join the coalition in the first stage.

Thirdly, in a model with heterogeneous players the number of players involved in the coalition is not always a good indicator of the ambition level of the coalition. In order to achieve high global abatement levels, it is more important to involve key players than to involve as many players as possible. For instance, the coalition of USA, Japan and the Rest of the World outperforms almost all other equilibria in the first stage.

Fourthly, in contrast to the one-shot game, China is not always present in the most ambitious equilibria. Coalitions that contain the USA with either Japan or EU15 perform better, but are not stable in the one-shot game. By opening up the possibilities for renegotiations these key players can be persuaded to join a coalition. Although they may be worse off in the short run, it improves the present value of their net benefits over the entire planning period.

Table 3. Best performing stable coalitions – renegotiations after 50 years

1 st stage	Coalition members		NPV of payoff [billion US\$]		
	1 st stage	2 nd stage	1 st stage	2 nd stage	Total
USA, EU15, EET, EEX, CHN, IND	USA, EET, CHN, IND, DAE	USA, EET, CHN, IND, DAE	6510	4409	10919
USA, EU15, EET, EEX, CHN, IND	EU15, EET, EEX, CHN, IND	EU15, EET, EEX, CHN, IND	6510	4402	10911
USA, JPN, ROW	USA, EET, CHN, IND, DAE	USA, EET, CHN, IND, DAE	6469	4409	10878
USA, EET, EEX, CHN, IND, DAE, ROW	USA, EET, CHN, IND, DAE	USA, EET, CHN, IND, DAE	6463	4409	10872
USA, JPN, ROW	EU15, EET, EEX, CHN, IND	EU15, EET, EEX, CHN, IND	6469	4402	10871
USA, OOE, EET, FSU, EEX, CHN, IND	USA, EET, CHN, IND, DAE	USA, EET, CHN, IND, DAE	6458	4409	10868
USA, EET, EEX, CHN, IND, DAE, ROW	EU15, EET, EEX, CHN, IND	EU15, EET, EEX, CHN, IND	6463	4402	10865
USA, EET, FSU, EEX, CHN, IND, DAE	USA, EET, CHN, IND, DAE	USA, EET, CHN, IND, DAE	6452	4409	10861
USA, OOE, EET, FSU, EEX, CHN, IND	EU15, EET, EEX, CHN, IND	EU15, EET, EEX, CHN, IND	6458	4402	10860
USA, EET, FSU, EEX, CHN, IND, DAE	EU15, EET, EEX, CHN, IND	EU15, EET, EEX, CHN, IND	6452	4402	10854

Table 4 gives the main regional results for the best performing sequence of coalitions. As there are no punishments in the second (last) stage by definition, the equilibria in this stage are not only self-enforcing, but also individually beneficial: all players are worse off when they leave the coalition, though the differences are sometimes small. This can be seen from the negative incentives to change membership in the second stage.

The punishment options are clearly visible in the incentives to change membership for the coalition members in the first stage. All coalition members would be better off if they could free-ride and leave the coalition. Their decision is, however, not based on a simple comparison of net benefits within the first stage, but on their net benefits over the full century. As free-riding would be punished in the second stage and the punishments are sufficiently severe, it is possible to enforce cooperation in the first stage.

As in the one-shot game, Japan and EU15 benefit strongly from the coalition that is formed, as their climate damages are substantially reduced, while their own abatement costs are moderate.

Table 4. Key results for the coalition of {USA, EU15, EET, EEX, CHN, IND} and {USA, EET, CHN, IND, DAE} in the first and second phase, respectively – renegotiations after 50 years

	Abatement		NPV of payoff		NPV of transfer		Incentive to change signal (NPV)	
	2011	2061	1 st com- mitment period	2 nd com- mitment period	1 st com- mitment period	2 nd com- mitment period	1 st com- mitment period	2 nd com- mitment period
	(% of BaU emissions)	(% of BaU emissions)	Billion US\$	Billion US\$	Billion US\$	Billion US\$	Billion US\$	Billion US\$
USA	17.5	8.7	1183	726	-358	-291	148	-4
JPN	2.5	2.6	1426	865	0	0	-247	-91
EU15	13.8	5.7	1293	1161	-514	0	162	-120
OOE	5.6	2.9	285	173	0	0	-39	-8
EET	37.1	17.3	90	62	68	28	11	0
FSU	6.7	5.0	554	336	0	0	-86	-26
EEX	20.8	1.7	203	150	77	0	26	-7
CHN	67.5	26.8	264	187	628	198	33	-1
IND	50.3	16.9	321	218	99	36	40	-1
DAE	1.9	13.0	206	117	0	29	-28	-1
BRA	0.1	0.1	127	77	0	0	-16	-2
ROW	6.3	4.4	558	338	0	0	-87	-26
Global			6510	4409	0	0		

The pattern of transfers across coalition members is conform expectations: the OECD members co-finance the abatement of greenhouse gases in non-OECD regions. Thus, the regions with large marginal benefits benefit from the high abatement levels in the regions with low marginal abatement costs. Coalitional maximisation requires that marginal abatement

costs are equal across coalition members and equal to the sum of marginal benefits of all coalition members. For singletons, marginal abatement costs equal their own marginal benefits.

Optimal renegotiation time

Earlier renegotiations may induce even larger coalitions in the first stage, as second stage payoffs are large compared to first stage payoffs. Thus, the threats that can be imposed on regions in the first stage are relatively large. As it is likely that the first stage equilibria are more ambitious than the equilibria in the second stage (see also Weikard, Dellink and van Ierland, 2006) and the first stage is now relatively short, this may not increase the performance of the resulting equilibria over the century. It may, however, boost cooperation in the first stage. While earlier renegotiations thus increase threats that help to stabilise larger coalitions, later renegotiations maintain the benefits of a larger coalition for longer. As these two mechanisms counteract each other, there is an optimal renegotiation moment where the additional benefit from high ambition balances with prolonged ambition in the first phase.

We find that the moment of renegotiation only affects the equilibria that emerge in the first commitment period, as shown in Table 5. The RPCs in the second commitment period do not depend on the length of the period, indicating that the self-enforcing equilibria that emerge without the possibility for punishment are robust with respect to the moment of renegotiation. In the first commitment period, however, the possibilities for credible punishment are crucial in the stabilisation of more ambitious coalitions: the later the renegotiations, the lower the threats and hence the less ambitious the best performing RPC.

Table 5. Key results for different moments of renegotiations

Renegotiations after ... years	# of RPC	Best performing RPC			
		Members 1 st commitment period	Members 2 nd commitment period	NPV of payoff	Concentrations 2110
10	35981	USA, JPN, EEC, OOE, EET, FSU, EEX, CHN, IND, DAE, BRA, ROW	USA, EET, CHN, IND, DAE	10451	1399
20	24256	USA, JPN, EEC, OOE, EET, FSU, EEX, CHN, IND, DAE, BRA, ROW	USA, EET, CHN, IND, DAE	11073	1393
30	35333	USA, EEC, OOE, EET, FSU, CHN, IND, DAE, ROW	USA, EET, CHN, IND, DAE	11308	1393
40	42949	USA, OOE, EET, FSU, EEX, CHN, IND, ROW	USA, EET, CHN, IND, DAE	11175	1395
50	35671	USA, EEC, EET, EEX, CHN, IND	USA, EET, CHN, IND, DAE	10919	1394
60	28668	USA, EET, EEX, CHN, IND, ROW	USA, EET, CHN, IND, DAE	10647	1397
70	27571	USA, OOE, EET, EEX, CHN, IND	USA, EET, CHN, IND, DAE	10397	1399
80	25925	EEC, EET, CHN, IND, ROW	USA, EET, CHN, IND, DAE	10096	1400
90	18858	USA, EET, EEX, CHN, IND	USA, EET, CHN, IND, DAE	9938	1402

When the renegotiation moment is sufficiently early, the Grand Coalition can be stabilised in the first commitment period. In fact, most large coalitions are equilibria when renegotiations are held after 10 or 20 years, including the 11-player coalitions where Brazil or Japan free-ride. The intuition behind this result is that these two regions have very steep marginal abatement cost curves. Their participation in the coalition requires that a substantial share of the surplus generated by the coalition would go to these regions to compensate their high abatement costs, but their participation would hardly benefit other regions, as their abatement levels would remain relatively small.

Because the larger stable coalitions exist for a shorter first commitment period, they do not contribute very much to the payoff over the entire planning horizon. Though the Grand Coalition is stabilised when renegotiations are held after 20 years, only 59% of the total potential gains of cooperation are reaped (100% in the first stage, 46% in the second stage). The best moment to renegotiate, in terms of maximising payoff of the best performing RPC, is after 30 years: this balances a sufficiently long first commitment period to enable prolonged strong abatement policies with sufficiently high threats to ensure stability of this ambitious coalition in the first commitment period; in this case, 61% of the potential gains of cooperation are reaped (in contrast, the worst renegotiation moment is after 90 years, where only 47% of the potential gains may be reaped – hardly better than can be achieved in the one-shot game).

Table 6. Key results for the coalition of {USA, EEC, OOE, EET, FSU, CHN, IND, DAE, ROW} and {USA, EET, CHN, IND, DAE} in the first and second phase, respectively – renegotiations after 30 years

	Abatement		NPV of payoff		NPV of transfer		Incentive to change signal (NPV)	
	2011	2041	1 st com- mitment period	2 nd com- mitment period	1 st com- mitment period	2 nd com- mitment period	1 st com- mitment period	2 nd com- mitment period
	(% of BaU emissions)	(% of BaU emissions)	Billion US\$	Billion US\$	Billion US\$	Billion US\$	Billion US\$	Billion US\$
USA	20.0	10.6	898	1071	-318	-434	229	-6
JPN	2.5	2.8	1149	1282	0	0	-267	-135
EU15	15.9	6.5	985	1721	-459	0	251	-179
OOE	26.5	3.8	173	256	15	0	44	-12
EET	41.9	20.7	66	92	54	41	17	-1
FSU	23.2	5.3	320	497	8	0	82	-39
EEX	1.9	1.8	200	223	0	0	-43	-10
CHN	77.5	30.6	205	277	567	296	52	-2
IND	57.4	21.3	236	322	72	54	60	-2
DAE	29.6	15.1	125	173	48	42	32	-1
BRA	0.1	0.1	102	114	0	0	-21	-2
ROW	26.9	4.9	322	501	12	0	82	-39
Global			4779	6529	0	0		

The regional results of the overall best performing RPC are presented in Table 6. The most striking result is that all signatories have an incentive to leave the coalition in the first stage, if they would not be punished in the second stage. For some regions, abatement levels are very large in the first stage, especially for China and India. In the absence of transfers, China would lose more than 350 billion US\$ on their own abatement efforts³. Clearly, they need to be compensated for this with substantial transfers. The major OECD signatories, USA and EU15, are willing to provide these transfers, as their benefits from global abatement are very high. In absolute terms (NPV of payoff), they are still the main beneficiaries of the agreement, together with Japan, who benefits from freeriding. Thus, the international flow of funds partially equalises payoffs across signatories.

Figure 1 shows annual global abatement percentages for the best performing RPC of the different specifications and, for comparison, for the All-Singletons case. In all cases, abatement percentages are falling for the first 7 decades, but increase slightly thereafter. The falling percentages are caused by increases in baseline emission levels that have a bigger impact than the increased abatement levels induced by technological progress. The slowdown of emission growth in the later decades causes abatement percentages to slightly increase towards the end of the century (cf. Nagashima *et al.*, 2005). The figure clearly shows the trade-off between ambition level and duration of the first commitment period, and also shows that the optimal transfer scheme is able to induce substantially higher abatement levels throughout the century than when no agreement is signed.

³ This can be seen in the table as the payoff minus the transfers received.

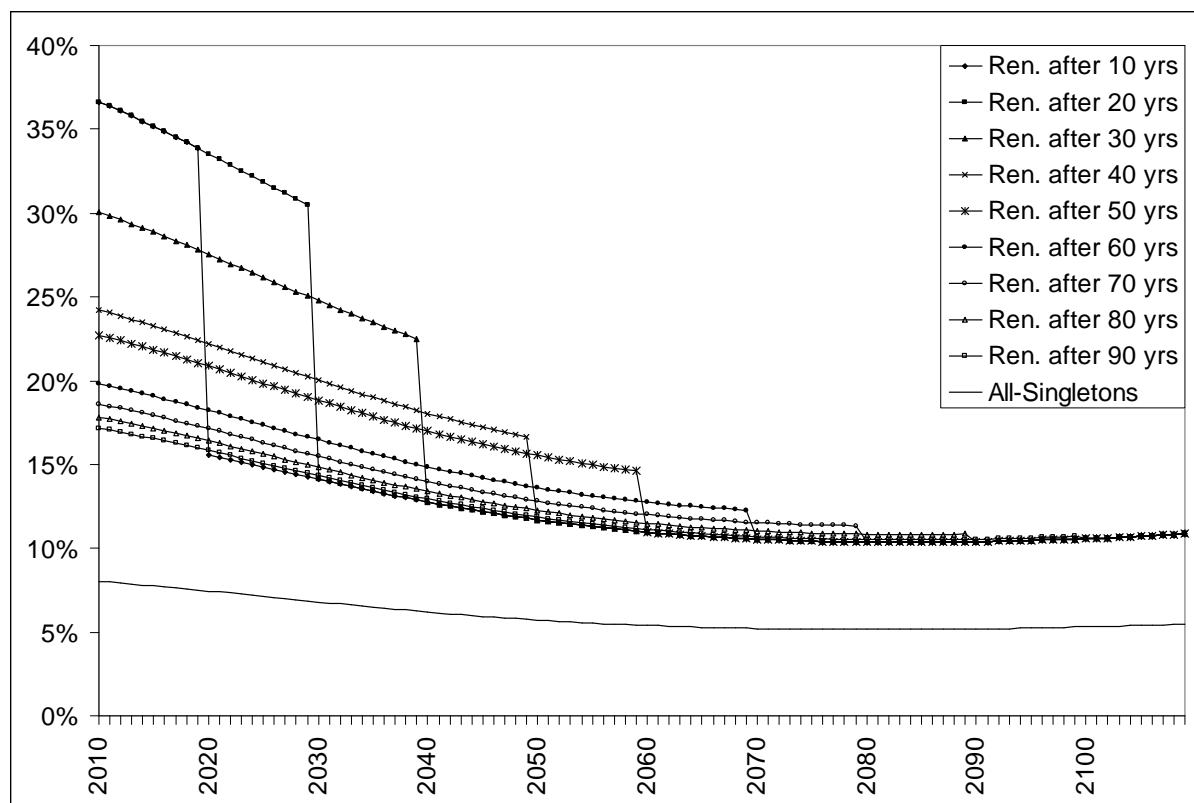


Figure 1. Global emission reductions in percentages for the All-Singletons case and the best performing RPCs in the different simulations

Conclusion

Although our analysis employs a cautious equilibrium concept – we assume that a region would defect from a coalition if its free-rider payoff is larger than its payoff in the coalition – we find equilibria with a large degree of cooperation. Our findings are in contrast to many other studies that find only small stable coalitions consisting of no more than two or three players; see *e.g.* the Barrett's (1994) theoretical analysis and the analysis of Finus *et al.* (2005) using the STACO model. There are three main drivers of our results. First, a well-designed transfer scheme is a useful tool to stabilise larger coalitions, as in our setting players differ in marginal benefits and marginal cost. Second, the transfer scheme that we suggest is specifically designed to increase the incentives to join the coalition. With such optimal transfers internal stability can be achieved for the largest possible set of coalitions. Third, we introduce renegotiations. A second commitment period can be used to threaten potential free-riders into cooperation.

References

- Altamirano-Cabrera, Juan-Carlos / Finus, Michael (2006) Permit Trading and Stability of International Climate Agreements. *Journal of Applied Economics* 9, 19-47.
- Babiker, M.H., Reilly, J.M., Mayer, M., Eckaus, R.S., Wing, I.S., Hyman, R.C. (2001) The MIT emissions prediction and policy analysis (EPPA) model: revisions, sensitivities and comparison of results. MIT joint program on the science and policy of global change. Report No.71, Cambridge.
- Barrett, Scott (1994) Self-enforcing international environmental agreements. *Oxford Economic Papers* 46, 878-894.
- Bernheim, B.D. / Peleg, B. / Whinston, M.D. (1987) Coalition-Proof Nash Equilibria I. Concepts. *Journal of Economic Theory* 42, 1-12.
- Bernheim, B.D. / Ray, D. (1989) Collective Dynamic Consistency in Repeated Games. *Games and Economic Behavior* 1, 295-326.
- Bosello, Francesco / Buchner, Barbara / Carraro, Carlo (2003) Equity, Development and Climate Change Control. *Journal of the European Economic Association* 1(2-3), 600-611.
- Botteon, Michele and Carraro, Carlo (1997) Burden-Sharing and Coalition Stability in Environmental Negotiations with Asymmetric Countries. In: Carraro, Carlo (ed.) *Environmental Policy, International Agreements and international Trade*. Cheltenham: Edward Elgar. 26-55.
- Carraro, Carlo / Siniscalco, Domenico (1993) Strategies for international protection of the environment. *Journal of Public Economics* 52, 309-328.
- Chander, Parkash / Tulkens, Henry (1995) A Core-theoretic Solution for the Design of Cooperative Agreements on Transfrontier Pollution. *International Tax and Public Finance* 2, 279-293.
- Dasgupta, Partha (1982) *The control of resources*. Oxford: Blackwell.
- d'Aspremont, Claude / Jaquemin, Alexis / Gabszewicz, Jean Jaskold / Weymark, John A. (1983) On the stability of collusive price leadership. *Canadian Journal of Economics* 16(1), 17-25.
- Dellink, R., Altamirano-Cabrera, J.C., Finus, M., van Ierland, E., Ruijs, A., Weikard, H.P. (2004) Empirical background paper of the STACO model. Mimeo, Wageningen University.
- Ellerman, A.D. / Decaux, A. (1998) Analysis of Post-Kyoto CO2 Emissions Trading Using Marginal Abatement Curves. MIT Report #40. Cambridge, Mass.: MIT.
- Eyckmans, Johan (1999) Strategy Proof Uniform Effort Sharing Schemes For Transfrontier Pollution Problems. *Environmental and Resource Economics* 14(2), 165-189.
- Eyckmans, Johan / Finus, Michael (2004) An Almost Ideal Sharing Scheme for Coalition Games with Externalities. Katholieke Universiteit Leuven, Center for Economic Studies, Working Paper 2004-14.
- Farrell, Joseph / Maskin, Eric (1989) Renegotiation in Repeated Games. *Games and Economic Behavior* 1, 327-360.
- Finus, Michael / Rundshagen, Bianca (2003) Endogenous coalition formation in global pollution control: a partition function approach. Carraro, Carlo (ed., 2003) *The Endogenous Formation of Economic Coalitions*. Cheltenham: Edward Elgar. 199-243.
- Finus, Michael / van Ierland, Ekko / Dellink, Rob (2006) Stability of Climate Coalitions in a Cartel Formation Game. *Economics of Governance* 7, 271-291.
- Hoel, Michael (1992) International environmental conventions: the case of uniform reductions of emissions. *Environmental and Resource Economics* 2, 141-159.
- McGinty, Matthew (2007) International environmental agreements among asymmetric nations. *Oxford Economic Papers* 59, 45-62.
- Nagashima, Miyuki / Dellink, Rob / van Ierland, Ekko (2006) Dynamic transfer schemes and stability of international climate coalitions. Mansholt Discussion Paper No. 23. Wageningen: Mansholt.
- Nordhaus, William D. / Boyer, J. (2001) *Warming the world: economic models of global warming*. Cambridge: MIT Press.

- Rose, Adam / Stevens, Brandt / Edmonds, Jae / Wise, Marshall (1998) International Equity and Differentiation in Global Warming Policy. An Application to Tradeable Emissions Permits. *Environmental and Resource Economics* 12, 25-51.
- Rubio, Santiago J. / Ulph, Alistair (2007) An infinite horizon model of dynamic membership of international environmental agreements. *Journal of Environmental Economics and Management* 54, 296-310.
- Tol, Richard S.J. (1997) A decision-analytic treatise of the enhanced greenhouse effect. PhD-thesis, Amsterdam: Vrije Universiteit.
- Weikard, Hans-Peter (2005) Cartel Stability Under an Optimal Sharing Rule. FEEM Working Paper No.77.05. Venice: FEEM.
- Weikard, Hans-Peter, Dellink, Rob and Ekko van Ierland (2006a) "Renegotiations in the Greenhouse". Paper presented at the Third World Congress of Environmental and Resource Economists, July 3-7, 2006, Kyoto, Japan.
- Weikard, Hans-Peter / Finus, Michael / Altamirano-Cabrera, Juan Carlos (2006b) The impact of surplus sharing on the stability of international climate agreements. *Oxford Economic Papers* 58, 209-232.