Fondazione Eni Enrico Mattei

Negotiating Climate Change as a Social Situation

Wietze Lise*, Richard S.J. Tol** and
Bob van der Zwaan*

NOTA DI LAVORO 44.2001

JUNE 2001

COALITION THEORY NETWORK

*Institute for Environmental Studies, Vrije Universiteit, Amsterdam
**Centre for Marine and Climate Research, Hamburg University, Institute for Environmental Studies, Vrije Universiteit, Amsterdam and
Center for Integrated Study of the Human Dimensions of Global Change, Carnegie Mellon University, Pittsburgh

Fondazione Eni Enrico Mattei
Corso Magenta, 63, 20123 Milano, tel. +39/02/52036934 – fax +39/02/52036946
E-mail: letter@feem.it
C.F. 97080600154

Negotiating Climate Change as a Social Situation

Wietze Lise[a], Richard S. J. Tol[b,a,c], and Bob van der Zwaan[a]

[a] *Institute for Environmental Studies, Vrije Universiteit, Amsterdam, The Netherlands*

[b] *Centre for Marine and Climate Research, Hamburg University, Hamburg, Germany*

[c] *Center for Integrated Study of the Human Dimensions of Global Change, Carnegie Mellon University, Pittsburgh, PA, USA*

DRAFT: June 21, 2001

Abstract

This paper applies the theory of social situations to international environmental agreements on greenhouse gas emission reduction. The usual pessimism on the size of stable coalitions among world regions is challenged for two alternative cases, namely by introducing farsightedness and by introducing coalitional moves with commitment. This is an extension of stability in the cartel game, where a cartel symbolises a coalition among world regions for reducing greenhouse gas emissions. It is a special case of the commitment situation, which has been proposed in the theory of social situations. The results are obtained by restricting the move rules in the game among world regions.

Keywords: Coalitions, coalitional moves, cooperation, theory of social situations, international negotiations, climate change.
JEL-classification: C7, F42, Q2

Non-technical summary

This paper studies the likelihood of coalitions to emerge between world regions which are negotiating over international environmental agreements on greenhouse gas emission reduction. The likelihood of international environmental agreements to be accepted increases in two alternative cases, namely for world region whom are more forward looking and for world regions whom are willing to commit to a certain greenhouse gas emission reduction target.

1. Introduction

Cooperation between countries is required for a substantial reduction of greenhouse gas emissions. However, in the absence of a world government it is difficult to create binding agreements; only self-enforceable agreements are likely. Greenhouse gas emission reduction is not only an international public good, but meaningful emission abatement requires a long-term, economy-wide effort. This sets climate change apart from the depletion of the ozone layer, where a technical fix, affecting only a small part of industry for a short time, solved the problem. This paper explores, from a game theoretic point of view, the scope for cooperation for greenhouse gas emission reduction. The game theoretic literature on climate change and other international environmental problems is large and growing fast. We use a more general approach than is typically the case.

An increase in the atmospheric concentration of greenhouse gases leads to global warming. Without emission abatement, atmospheric concentrations would double in the second half of the 21st century, predicted to cause a global average temperature increase of 1.5–6 °C and a rise in sea level of some 20-80 centimetre by 2100 (IPCC, 2000). This could have a number of negative effects, including increased risks for river floods, an extension of malarial areas, and the disappearance of small islands , although there are also positive effects, particularly on agriculture (Smith et al., forthcoming).

The international community takes the problem of global warming serious, as witnessed by the United Nations Framework Convention on Climate Change and its Kyoto Protocol. This protocol sets a target for GHG-emission abatement in 2012 for developed countries and economies in transition. The target has been set, but it will not become effective until it is ratified by a group of countries that contribute collectively at least 55% to the GHG-emissions. So far, only a few small countries have ratified the protocol whereas it is increasingly unlikely that key emitters such as the United States and the European Union will implement the agreement.

Many authors characterise the problem of reducing greenhouse gas emission as a prisoner's dilemma (Barrett, 1994, 1998a, 1998b, 1998c; Carraro 1997a, 1997b, 1997c, 1998, 1999; Cesar, 1994; Ecchia and Mariotti, 1998, among others), reflecting the fact that emissions mix uniformly in the atmosphere. This implies that each country benefits from greenhouse gas emission reduction by other nations, but has little incentive to do so itself.

Besides global cooperation in a "grand coalition", intermediary sized coalitions are also important. A simple but effective way for doing so is by following the theory of cartel stability (d'Aspremont and Gabszewicz, 1986). An advantage of this framework is that it has a fairly large set of unique solutions. Cartel stability has three characteristics. It analyses whether a world region in the cartel has no incentive to leave the cartel (internal stability), whether a world region outside the cartel has no incentive to join the cartel (external stability) and whether all members of the stable coalition are better off as compared to the case without any coalition (profitability). A coalition is called stable when it is both internally and externally stable.

The cartel game can be extended in many ways. One possibility is the case where two world regions decide simultaneously to enter or leave the cartel. We refer to such decisions as coalitional moves. This extension is useful as many practical situations can be represented by coalitional moves. Another extension of cartel stability is farsighted behaviour (Chwe, 1994), where a deviation of one world region can trigger a deviation by another world region, and so on. This paper elaborates on these two extensions by using the theory of social situations (TOSS) proposed by Greenberg (1990). TOSS is flexible in incorporating coalitional moves and farsighted behaviour explicitly.

The TOSS differs from a standard game theoretic approach in many ways. In a TOSS game players can move from one position to another. A position in the game is characterised by the set of players, the set of possible outcomes which can be obtained from that position onwards, and a utility function which attaches a value to these outcomes. The inducement correspondence specifies the rules of move from one position to another (e.g., from cooperative to non-cooperative behaviour) in the game. A TOSS game is completed by a standard of behaviour (SB), which is suggested to the players of the game. This SB tells the players which outcomes in each position are solutions of the game. The solution for a TOSS game exists and is unique once it is hierarchical. Many games of practical interest, including the games of this paper, have this property.

In this paper we also extend the cartel game. Of course, many other extensions of the cartel game can be thought of. For instance, Ray and Vohra (1997) developed a technique to calculate the coalition structure of an *n*-person game, where a collection of coalitions can emerge simultaneously. They use a model of symmetric players with utility transfer to show that a (symmetric) coalition structure can emerge in such situations. They generate a recursive method for generating the coalition structure. Another example, Na and Shin (1998) introduce uncertainty in the payoffs and show that the likelihood of the grand coalition to be stable is greater the sooner climate change negotiations take place. Ulph and Maddison (1997) essentially conclude the same thing, and argue that learning may thus decrease the possibilities for cooperation.

This paper is organised as follows. Section 2 gives an overview of recent game theoretic approaches to climate change negotiations. Where relevant, we point out the scope for applying game theoretic models to climate change negotiations. Section 3 sets up a TOSS game and points out its differences and commonalities with a non-cooperative game. We apply the theory of social situations to the cartel game in Section 4 by introducing an aspect of farsightedness in the game by allowing for two deviations. In Section 5, we set up an incremental individual and coalitional commitment situation with three world regions to study the possibilities of cooperation in climate change negotiations. We apply the commitment model to the situation where international trade matters, using empirical data from a general equilibrium model. We consider the scope for forming coalitions between 3 world regions. The final section points out the main conclusions and areas for further research.

2. Literature on games in climate change negotiations

The class of cooperative games constitutes an obvious set of games to study the stability of coalitions. Carraro (1997a, 1997b, 1998, 1999), Botteon and Carraro (1997, 1998), Carraro and Siniscalco (1997) exploit cooperative games, particularly the theory of cartel stability by exploring the likelihood of international environmental agreements to emerge. Carraro also studies the possibilities of enlarging the size of the stable coalition by considering transfers between world regions and issue linkage between emission reduction on the one hand and trade or technology on the other. Carraro's main objective is to analyse the maximum number of world regions in a stable coalition to see whether self-enforcing international environmental agreements are possible.

In contrast to Carraro, Barrett (1994, 1998a, 1998b, 1998c) follows a non-cooperative game theoretic approach in studying the stability of self-enforcing international environmental agreements (IEAs). His one-shot strategic game shows the difficulty in obtaining a coalition of a large size. He also explores the possibility of punishing free-riders through different means, such as trade sanctions and a temporal deviation of cooperating countries, and studies the credibility of the threat for deterring free-riding behaviour in IEAs. He concludes that repeated negotiations and the threat of sanctions can enlarge the size of the coalition, but the grand coalition, i.e. cooperation among all world regions, is hard to obtain.

Carraro and Siniscalco (1993), however, claim that repeated games are not a useful tool, because the observed behaviour in international negotiations has no resemblance with trigger strategies.[1] This is not only a practical objection, but also theoretical; trigger strategies in global warming result in self-punishment, as the region which implements the trigger strategy increases his emissions, which they are trying to reduce jointly. Hence, more complicated behaviour need to be studied.

An important result for repeated games is the Folk theorem, which states that 'any feasible individual rational payoffs can arise in an equilibrium if the players are sufficiently patient' (Fudenberg and Maskin, 1986) in a repeated game. More specifically, it is always possible to sustain a higher payoff than the minimax payoff when the discount factor is sufficiently large in the infinitely repeated game. The minimax payoff is the highest payoff a player can get when all other players switch to the most unfavourable action for this player. Fudenberg and Maskin (1986) claim that this result also holds for finitely repeated games with incomplete information.

[1] A trigger strategy consists of two options, where a player cooperates until he observes a certain undesired behaviour among other players in the game and the player switches to a competitive strategy, to "punish" the other players for not living up to his expectations.

Greenberg *et al* (1996) derive, under stringent conditions, existence and uniqueness in an extension of a multistage repeated game using the framework of social situations. Two fundamental assumptions within a repeated game theoretical approach are relaxed in their paper. First, the assumption that the strategy space is a Cartesian product is dropped. This allows for sequential moves, rather than the more commonly used simultaneous moves. Second, the subgame perfectness requirement (the need to consider each truncation of a path) is no longer necessary in social situations.

Alternatively, the stability of coalitions can also be studied by setting up a TOSS game. In a TOSS game, we no longer talk about equilibria, but about situations. Situations differ from equilibria in the sense that the move rules of the game are pointed out explicitly, while they are implicit in a conventional game. Greenberg (1990) describes four situations with coalitional moves. This are all extensions of the Nash equilibrium, by allowing a group of players to deviate simultaneously and/or allowing more than one deviation.

Strong Nash situation considers besides individual deviations (from cooperation) also group deviations. This requirement is demanding, as the solution set of many games of interest is empty. This is, for instance, the case for the prisoner’s dilemma.

Coalition-proof Nash situation allows for more than one deviation. The deviation sequence is finite by only allowing members of a previously deviating group of players to deviate further, and so on. The players outside the deviating group of players are not allowed to move. The probability of an empty solution set to occur is smaller than in the previous case.

Coalitional contingent threat situation is similar to the coalition-proof Nash situation. The group of deviating players, however, need no longer be a subset of the previous group of deviating players. This situation is fully farsighted. Obviously, some restrictions need to be applied on the deviation process in order to induce a non-empty solution set. For instance, infinite sequences of a group of deviating players need to be excluded.

Coalitional commitment situation: In this situation, step by step, a group of players can commit themselves to a certain strategy from which they can no longer deviate in future. This process continues until all players have committed themselves to a certain strategy. Under some mild conditions the solution set is non-empty. We consider this situation in Section 5.

These four situations have been used by a number of authors. DeMarzo (1992) builds on the strong Nash situation and the coalition-proof Nash situation. His model is an *n*-person, *m*-strategy repeated multistage game (which allows for changing payoffs), where coalitional moves are allowed and communication can take place. DeMarzo (1992) is not able to give a complete characterisation of the solution set and falls back on the 2-person case in order to find explicit results.

Chwe (1994), Mariotti (1997), Ecchia and Mariotti (1998) and Nakanishi (1999) build their argument on the coalitional contingent threat situation. Chwe’s (1994) main argument is to plead for farsighted behaviour, where players anticipate other player’s reasoning. The possible outcomes in a farsighted framework can be quite different, as we also show in Section 4. Eyckmans (2000) applies Chwe's work to climate change. Mariotti (1997) and Ecchia and Mariotti (1998) use the coalitional contingent threat situation to define a coalitional game. Under strong assumptions, they make a first characterisation about conditions for non-emptiness of the solution set, which they label as coalitional equilibrium. Nakanishi (1999) uses the contingent threat situation, but he omits coalitional moves. His intuitive application of TOSS shows that alternative standards of behaviour that can lead to cooperation in the international export quota retaliation game, while a Nash behaviour in such a game leads to the conclusion that trade barriers increase and international trade goes down to zero.

3. The theory of social situations

In general, the main objective of non-cooperative game theory is to verify for each possible strategy bundle whether or not it is an equilibrium. Such an equilibrium has an arbitrary institutional setting, which may or may not be representative for the desired application one wishes to study. A TOSS game, on the contrary, has a tool for explicitly formulating the institutional setting of the game. This leads to a quite different description of the game as compared to a conventional game. This section shows how a TOSS game can be formulated.

The building blocks of TOSS, namely positions, inducement correspondence and standards of behaviour differ from the conventional game theoretic approach so as to enable an explicit and formal description of the rules of the game.

The first building block is a position G, which constitutes a triplet (N, X^G, u), where

- N are the *players* of the game at position G,
- X^G is the set of possible *outcomes* of the game starting from position G,
- $u \equiv \{u_i\}_{i \in N}$, where utility function u_i assigns a value to all possible outcomes that can be attained by starting from position G, hence u_i: $X^G \rightarrow R$.

In general the number of players and the utilities can differ from position to position, but in this paper we only consider games where the set of players and the utilities of the outcomes for the players is the same in each position.

A position differs from a standard game in the sense that it considers the attainable outcomes in position G, instead of focusing on actions. The following example tries to clarify what is meant with a "position". The position in a 2-person game, where players can choose between cooperate (C) or defect (D) has the following shape:

$$\begin{aligned} G = (\{1,2\}, \{\text{CC,CD,DC,DD}\}, \{&u_1(\text{CC}), u_1(\text{CD}), u_1(\text{DC}), u_1(\text{DD}), \\ &u_2(\text{CC}), u_2(\text{CD}), u_2(\text{DC}), u_2(\text{DD})\}) \end{aligned} \tag{1}$$

The collection of all possible positions in a game is denoted by Γ.

The second building block of a social situation is the *inducement correspondence* (γ). The economic rationale of the inducement correspondence is to restrict a TOSS game by appropriate institutional rules for a given situation. Operator γ describes whether or not a coalition $s \subset N$ can move from current position G to alternative position $H \in$

$\gamma_s(G,x^*)$, when outcome x^* from the set of outcomes (X^G) is suggested. Each player of the game has to decide whether the proposed outcome x^* is acceptable or not. If a certain coalition *s* does not accept the proposed outcome x^*, they will *induce* alternative position *H* in case γ permits to do so. It is assumed that the members of the deviating coalition, while making the decision, are aware of the implications of their choice for the continuation of the game. In order to fully describe the game, it is required that all induced positions are an element of all positions in the game Γ. This means that all possible positions are well-defined in the game:

$$\text{For all } G \in \Gamma \text{ and s} \subset N \text{ and } x \in X^G : \gamma_s(G,x) \subset \Gamma \quad (2)$$

The following example illustrates how an inducement correspondence can be used to define the rules of a game. The inducement correspondence in a 2-person game based on the Nash equilibrium, where player's choices can be represented by the pair (x_1,x_2) and $x_i = \{C,D\}$ has the following shape:

$$\begin{aligned} &\gamma_i(G,x) = G^i, \text{ where } G^i = (\{i\},(x_i,\hat{x}_{-i}),u), \forall G \in \Gamma, i \in N, x \in X^G \\ &\gamma_s(G,x) = \varnothing, \text{ otherwise} \end{aligned} \quad (3)$$

We added a cap (^) above a variable to denote that it is fixed. Hence, in the 2-person Nash game, only individual moves are allowed and once a player has moved other players can make no additional moves.

Pair (γ, Γ) is called a situation, where γ describes the (move) rules of the game and Γ the set of attainable positions. Finding the most appropriate formulation of γ is the main challenge in setting up a social situation in a practical setting. In Sections 4 and 5 we work out explicit specifications of γ for climate change negotiations.

The third step in TOSS is to identify which outcomes (X^G) are *solutions* ($\sigma(G)$) in any position of the game. To this end, a so-called standard of behaviour is defined. The standard of behaviour is a norm imposed upon the players of the game, in order to find a solution in any position. In order to find meaningful solutions, it is required that the standard of behaviour is (internally and externally) stable.[2] Finally, we require an algorithm which compares solutions in position *G* with solutions in the induced position *H*. Let us, therefore, take the optimistic stable standard of behaviour (OSSB) as the guiding norm. The OSSB is internally stable when the utility of all outcomes *inside* the solution set are at least as good as the utility of an outcome in the solution set which can be induced by any deviating group of players. The OSSB is externally stable when the utility of all outcomes *outside* the solution set are strictly lower than the utility of at least one element in the solution set which can be induced by a deviating group of players. Hence, internal and external stability concerns the outcomes in- and outside the solution set, rather than members of a coalition. The OSSB is stable when it both internally and externally stable[3].

---

2 This is a deviation from a cooperative game, where outcomes require to be stable rather than the standard of behaviour.

3 Alternatively, the conservative standard of behaviour (CSSB) excludes an outcome as a solution if all members of a coalition *s* are better off with *all outcomes* in $\sigma(H)$. We prefer the optimistic standard of behaviour, as it gives a smaller number of outcomes in the solution set.

Greenberg designed a compact way of representing stability by setting up a set of optimistically dominated outcomes (ODOM(σ,G)). ODOM(σ,G) excludes only those outcomes $x \in X^G$ for which, in position $H \in \gamma_s(G,x)$, *all* members in coalition s are better off for *at least one* outcome in $\sigma(H)$. That is, the players move from a position G to a position H if it is possible, from H onwards, that they can reach a position with an outcome in which they are all better off. Formally, ODOM(σ,G) is defined as:

$$\begin{aligned}&\text{ODOM}(\sigma,G)=\{x\in X^G \mid \text{there exists } [s\subset N,\, H\in\gamma(s\mid G,x) \text{ and } y\in\sigma(H)]\\&\text{such that for all } i\in s: u_i(y)>u_i(x)\}\end{aligned} \tag{4}$$

The solution set in position G is defined as the complement of (4): $\sigma(G)=X^G\backslash\text{ODOM}(\sigma,G)$.

In the following sections, we try to find the elements of the standard of behaviour σ in any possible position. In all cases, we find a unique OSSB, but there is often more than one outcome in the solution set $\sigma(G)$.

4. The cartel situation

The cartel game stands at the basis of the stability analysis of international environmental agreements by Carraro (1997a,b, 1998, 1999). The cartel game can be described by two operators, namely $P_i(s)$ –the value for player i to be a member of coalition s– and $Q_i(s)$ –the value for player i not to be a member of coalition s. Table 1 illustrates this for the 3-person case in a normal form game. Since coalitions of one player make no sense, four cells of the payoff matrix have the same payoffs (Q_1(ø), Q_2(ø), Q_3(ø)).

Table 1 The cartel game as a normal form game.

3	1	2	Cooperate	Defect
Cooperate	Cooperate		$P_1(1,2,3)$, $P_2(1,2,3)$, $P_3(1,2,3)$	$P_1(1,3)$, $Q_2(1,3)$, $P_3(1,3)$
	Defect		$Q_1(2,3)$, $P_2(2,3)$, $P_3(2,3)$	Q_1(ø), Q_2(ø), Q_3(ø)
	1	2	Cooperate	Defect
Defect	Cooperate		$P_1(1,2)$, $P_2(1,2)$, $Q_3(1,2)$	Q_1(ø), Q_2(ø), Q_3(ø)
	Defect		Q_1(ø), Q_2(ø), Q_3(ø)	Q_1(ø), Q_2(ø), Q_3(ø)

A coalition is internally stable if $P_i(s) > Q_i(s\backslash i)$ for all $i \notin s$ – that is, no player wants to leave the coalition – and a coalition is externally stable if $P_i(s \cup i) < Q_i(s)$ for all $i \in s$ – that is, no player wants to join the coalition. The stability requirement of a cartel situation is equivalent to a Nash equilibrium if inequality rather than *strict* inequality signs are used.

The cartel situation consists of n regions ($N = \{1,2,\ldots,n\}$) in the world negotiating for greenhouse gas emission reduction. We denote the collective action as an n-tuple $(x_1,x_2,\ldots,x_n) \in X^N$, where $x_i \in \{C,D\} = X_i$ represents the choice of region i and $X^N = X_1 \times X_2 \times \ldots \times X_n$. This n-tuple can also be written as a pair (x_i,x_{-i}), where the minus sign for the i denotes all world regions except i. We add a cap (^) above a variable to denote that it is fixed. As the climate change negotiations have often been modelled as an n-person prisoner's dilemma, we do the same. The following inequality must hold in the case of a 'strict' prisoner's dilemma:

$$u_i\left(\mathrm{D},\hat{x}_{-i}\right) > u_i\left(\mathrm{C},\hat{x}_{-i}\right), \text{for all } i\in N \text{ and } G\in\Gamma \qquad (5)$$

This inequality shows that it is always better for a region to defect irrespective of the number of other regions who are cooperating. Hence, this payoff order has the characteristic that the n times D outcome ($\underbrace{\mathrm{D},\mathrm{D},\ldots,\mathrm{D}}_{n}$) is the single pure strategy Nash equilibrium.[4]

The cartel situation consists of just two positions, namely the initial position where all world regions can still move and a final position where just one world region can move. This reflects the mechanism of a Nash equilibrium that considers only one move of one region at a time, after which other regions are assumed not to move any further.

$G^N = \left(N, X^N, u\right)$

$G^i = \left(\{i\}, \left(x_i, \hat{x}_{-i}\right), u\right)$, where $\Gamma = G^N \cup G^i$

The inducement correspondence has the following shape:

$$\begin{array}{l}\gamma_i(G,x) = G^i, \forall G\in\Gamma, i\in N, x\in X^N \\ \gamma_s(G,x) = \varnothing, |s|>1\end{array}, \text{ hence } G^N \xrightarrow{\{i\}} G^i \circlearrowright.$$

The set of outcomes of this game as suggested by the OSSB is derived in two steps by eliminating the dominated outcomes (ODOM), viz.:

Step 1: Position G^i has two outcomes, namely $\left(\mathrm{D},\hat{x}_{-i}\right)$ and $\left(\mathrm{C},\hat{x}_{-i}\right)$. Since region i can keep on inducing G^i forever, we are able to apply equation (4). Because of our assumption in equation (5) we see that ODOM(σ,G^i) = $(\mathrm{C},\hat{x}_{-i})$ and hence $\sigma(G^i)$ = $(\mathrm{D},\hat{x}_{-i})$.

Step 2: Position G^N has 2^N outcomes. If an outcome is proposed in which a region has to choose 'C', then this outcome is rejected, because of equation (5). This holds for all $x\in X^N$ and $i\in N$, hence $\sigma(G^N)$ = ($\underbrace{\mathrm{D},\mathrm{D},\ldots,\mathrm{D}}_{n\text{ times}}$) is then unique outcome in the solution set.[5]

The optimal behaviour in the final position (G^i) is to choose the best reply, which is to defect. This shows that a TOSS game suggests how to behave outside of equilibrium. Interestingly, the outcome, resulting from G^N, is the Nash equilibrium, as only the outcome where all regions defect is suggested to the players by OSSB.

The cartel situation, however, is myopic, as players are not allowed to move after one deviation. There is no reason for world regions to behave in such a way. Let us consider the situation where after one deviation another region can deviate as well,

---

[4] If there are more than 2 players, other weak prisoner's dilemmas exist as well, where besides the "nobody cooperates" Nash equilibrium, other Nash equilibria exist as well.

[5] Since in both positions, the number of outcomes in the solution set is one, the two standards of behaviour OSSB and CSSB suggest the same course of action.

after which the game stops. Here, we are dealing with the case where the world regions are more forward looking, by allowing for one more deviation. This limited foresight cartel situation can be characterised by the following positions:

$$G^N = \left(N, X^N, u\right)$$
$$G^j = \left(N, \left(x_j, \hat{x}_{-j}\right), u\right)$$
$$G^{ij} = \left(\{i\}, \left(x_i, \hat{x}_{-i}\right), u\right), i \neq j, \text{ where } \Gamma = G^N \cup G^j \cup G^{ij}$$

The inducement correspondence has the following shape:

$\gamma_j(G^N, x) = G^j; j \in N, x \in X^N$

$\gamma_i(G, x) = G^{ij};\ \forall G \in G^j \cup G^{ij}; i, j \in N, i \neq j, x = (\hat{x}_j, x_{-j}) \in X^N$, hence

$\gamma_s(G, x) = \varnothing, |s| > 1, G \in \Gamma, x \in X^N$

$G^N \xrightarrow{\{j\}} G^i \xrightarrow{\{i\}} G^{ij}$.

The standard of behaviour (OSSB) for the limited foresight cartel situation can be derived in three steps as follows.

Step 1: the solution set associated with the final position G^{ij} is exactly the same as in the (myopic) cartel situation, because of the choice possibilities are described by the inducement correspondences and the final position are identical. Hence, $\sigma(G^{ij})$ = $(\text{D}, \hat{x}_{-i})$.

Step 2: In position G^j there are 2^N outcomes possible. In order to verify which outcomes are dominated by elements of the solution set in the induced position G^{ij}, we have to evaluate the following inequality: $u_i\left(\text{D}, \hat{x}_{-i}\right) > u_i\left(x_i, \hat{x}_{-i}\right)$. First of all, note that on both sides of the inequality the x_j (as an element of $\hat{x}_{-i}$) is fixed by inducing the final position G^{ij}. Hence, it does not matter whether x_j equals 'C' or 'D'. Then ODOM(σ,G^j) contains all positions where at least one region cooperates (x_i = 'C') after the choice of region j. This results in ODOM(σ,G^j) = $X^N/(\underbrace{\text{D,D,}\ldots\text{,D}}_{n \text{ times}})$ and

$\sigma(G^j) = \underbrace{\text{D,D,}\ldots\text{,D}}_{n \text{ times}}$.

Step 3: In position G^N, there are 2^N outcomes possible. As there is only one element in the solution set of the inducible position G^j, the following inequality has to be evaluated: $u_i(\underbrace{\text{D,D,}\ldots\text{,D}}_{n \text{ times}}) > u_i\left(x_i, \hat{x}_{-i}\right)$. It is easy to see that only one type of outcome is dominated, namely the outcome where region j cooperates and all other regions defect. Then the solution set is the complement and consists of 2^{N-1} outcomes.

We see that the threat of a further deviation deters the first region to deviate. Obviously, a further deviation is only possible if at least one of the other countries is cooperating in the suggested outcome. Hence, by including more foresighted behaviour and by adjusting the move rules of the world regions, we see that the grand coalition is no longer excluded from the solution set. Note that the number of outcomes in the solution set is far from unique, but the OSSB is. The solution set consists of all possible coalitions excluding the (trivial) the case where only one region cooperates. There is no reason to assume that this situation is less plausible for regions in the world than the myopic cartel situation. In the next sections, we explore other inducement correspondences to verify whether other solutions can be suggested.

5. Commitment

5.1 Individual incremental commitment situation

We consider the case of negotiations between three symmetric world regions. Let us consider the following simple set of institutional rules, where a single region has the choice between committing (C) or defecting (D). Here commitment is stronger than cooperation, as a committed region is no longer able to revise his strategy, while the uncommitted region is. For the moment, we focus on individual moves alone. We refer to this situation as the individual incremental commitment situation (IICS), which is a simplified version of the individual commitment situation (Greenberg, 1990, page 106-109). In the next subsection, coalitional moves are added to the model.

Since the three regions can choose between commitment or not, DDD, CDD, DCD, DDC, CCD, CDC, DCC and CCC are the 8 possible outcomes in the IICS. Climate change negotiations have the flavour of a prisoner's dilemma. This can be defined with the following inequality, where the first letter denotes the actions of player *i* and the following number denotes the number of committed world regions:

$$u_i(\mathrm{D2}) > u_i(\mathrm{C2}) > u_i(\mathrm{D1}) > u_i(\mathrm{C1}) > u_i(\mathrm{D0}) > u_i(\mathrm{C0}) \quad (6)$$

Table 2 shows the payoff matrix of this three-region game, satisfying equation (6). The utilities in this payoff matrix represent the payoff order.

Table 2 Payoff matrix for a 3-person symmetric prisoner's dilemma.

3:	1:	2: Commit	Defect
Commit	Commit	5,5,5	3,6,3
	Defect	6,3,3	4,4,1
	1:	2: Commit	Defect
Defect	Commit	3,3,6	1,4,4
	Defect	4,1,4	2,2,2

From this payoff matrix, we can see that the grey shaded outcome DDD is the Nash equilibrium. Would an explicit description of the negotiation process results in alternative outcomes of the game?

We can identify 8 positions in this situation; see Table 3.

Table 3 Positions in the individual incremental commitment situation.

G^0	({1,2,3},{CCC},u)
G^1	({1,2,3},{DCC,CCC},u)
G^2	({1,2,3},{CDC,CCC},u)
G^3	({1,2,3},{CCD,CCC},u)
G^{12}	({1,2,3},{DDC,DCC,CDC,CCC},u)
G^{13}	({1,2,3},{DCD,DCC,CCD,CCC},u)
G^{23}	({1,2,3},{CDD,CDC,CCD,CCC},u)
G^{123}	({1,2,3},{DDD,CDD,DCD,DDC,CCD,CDC,DCC,CCC},u)

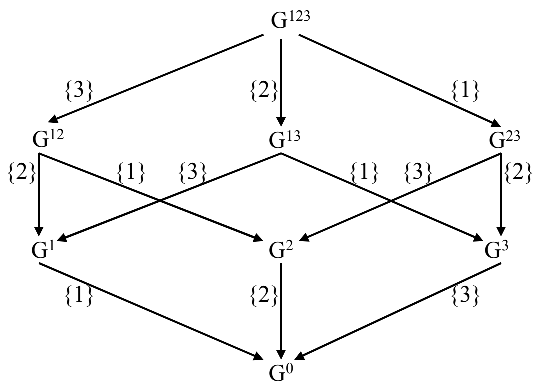
Table 4 shows the permitted moves by the players as expressed by the inducement correspondence for this case:

Table 4 Inducement correspondence for the individual incremental commitment situation.

$\gamma_1(G^1,\text{DCC}) = \gamma_2(G^2,\text{CDC}) = \gamma_3(G^3,\text{CCD}) = G^0$
$\gamma_1(G^{13},\text{DCD}) = G^3$; $\gamma_2(G^{23},\text{CDD}) = G^3$
$\gamma_1(G^{12},\text{DDC}) = G^2$; $\gamma_3(G^{23},\text{CDD}) = G^2$
$\gamma_2(G^{12},\text{DDC}) = G^1$; $\gamma_3(G^{13},\text{DCD}) = G^1$
$\gamma_1(G^{123},\text{DDD}) = G^{23}$; $\gamma_2(G^{123},\text{DDD}) = G^{13}$; $\gamma_3(G^{123},\text{DDD}) = G^{12}$
$\gamma_s(G,x) = \emptyset$, otherwise for $G\in\Gamma$; $x\in X^G$

Figure 1 depicts the inducement correspondence graphically.

Figure 1 The individual incremental commitment situation for three regions.



Because of the recursiveness in the definition of OSSB – cf. equation (4) – the outcomes in the solution set are computed by starting with a position which can only induce itself, which in this case is G^0. The ODOM is empty for each position, as a particular player can only induce a new position by switching from D to C, which decreases the utility for this particular player. This means that $\sigma(G)=X^G$ for all $G\in\Gamma$.

We see that the grand coalition is not excluded as an outcome in the solution set. We also observe that the 'Nash equilibrium' (DDD) can only be attained when initially all regions defect. This trivial result is not too helpful, as the unique OSSB suggests all possible outcomes belong to the solution set.

5.2 Coalitional incremental commitment situation

We now focus on coalitional moves. We refer to this situation as the coalitional incremental commitment situation (CICS). By allowing for coalitions the positions of the game will not change, but the possible moves do. Figure 2 shows how the inducement correspondence has to be extended to account for coalitional moves.

Figure 2 The coalitional incremental commitment situation for three regions.

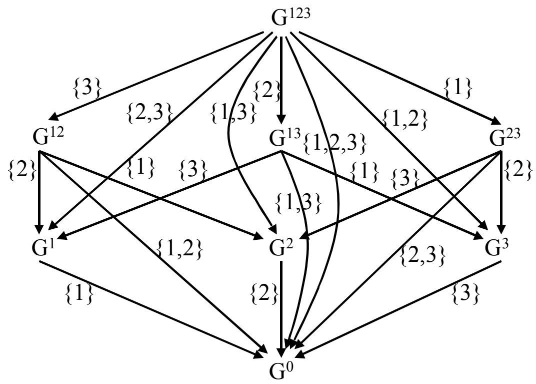


Table 5 shows the additions to the inducement correspondence in Table 4.

Table 5 Additional moves in the coalitional incremental commitment situation.

$\gamma_{\{1,2\}}(G^{12},\text{DDC}) = \gamma_{\{1,3\}}(G^{13},\text{DCD}) = \gamma_{\{2,3\}}(G^{23},\text{CDD}) = \gamma_{\{1,2,3\}}(G^{123},\text{DDD}) = G^0$
$\gamma_{\{1,2\}}(G^{123},\text{DDD}) = G^3$; $\gamma_{\{1,3\}}(G^{123},\text{DDD}) = G^2$; $\gamma_{\{2,3\}}(G^{123},\text{DDD}) = G^1$
$\gamma_s(G,x) = \emptyset$, otherwise for $G \in \Gamma$; $x \in X^G$

Table 6 shows the outcomes that are excluded from the solution set if coalitional moves are allowed for. The outcomes {DDD,CDD, DCD, DDC} do not belong to the solution set associated with position G^{123}, because a group of regions can improve their utility by inducing another position.

Table 6 The ODOM for the coalitional incremental commitment situation.

$\text{ODOM}(\sigma,G^0) = \emptyset$	$\text{ODOM}(\sigma,G^{12}) = \{\text{DDC}\}$
$\text{ODOM}(\sigma,G^1) = \emptyset$	$\text{ODOM}(\sigma,G^{13}) = \{\text{DCD}\}$
$\text{ODOM}(\sigma,G^2) = \emptyset$	$\text{ODOM}(\sigma,G^{23}) = \{\text{CDD}\}$
$\text{ODOM}(\sigma,G^3) = \emptyset$	$\text{ODOM}(\sigma,G^{123}) = \{\text{DDD,CDD,DCD,DDC}\}$

The likelihood of cooperation is considerably higher in the case of CICS than in the case of IICS. Under CICS, at least two players cooperate.

5.3 The trade situation: an application of the 3-person incremental coalitional commitment situation

We take the analysis one step further by analysing four world regions, the fourth one (the rest of the world) being a dummy player that continues its emissions unabated:

1. United States (US),
2. European Union (EU),
3. Japan (JAP), and
4. Rest of the world (ROW).

With this division, we can use data from a recursive-dynamic, multi-regional computable general equilibrium (Kemfert *et al*, 2001). As before, we consider the following simple set of institutional rules, where a single region has the choice between committing (C) or defecting (D). If at least two regions commit, they form a coalition and maximise joint welfare, otherwise they maximise their welfare unilaterally.

We allow coalitional moves and assume that once a region commits to cooperate, it can no longer revise their strategy. Since the three regions can choose between committing or defecting and at least two regions need to commit in order to form a coalition, DDD, CCD, CDC, DCC and CCC are the 5 possible outcomes in this situation.

Table 7 shows the estimated payoff order for this game, after Kemfert *et al.* (2001). As before, the utilities in this payoff matrix represent the payoff order. Three cases are distinguished. In the “No Trade” case, the only interaction between regions is the global climate. In the two trade cases, regions’ emission reduction costs depend on the other regions’ emission reduction. In the “Trade 1” case, only total emission reduction costs are affected by the other regions’ abatement. In the “Trade 2” case, both the total and the marginal costs are affected.

Table 7 Payoff matrices for the trade situation.

NoTrade:

JAP	US	EU	Commit	Defect
Commit	Commit		1,5,5	2,4,3
	Defect		5,1,2	3,2,1
	US	EU	Commit	Defect
Defect	Commit		4,3,4	3,2,1
	Defect		3,2,1	3,2,1

Trade1:

JAP	US	EU	Commit	Defect
Commit	Commit		5,4,5	4,5,3
	Defect		3,1,2	1,2,1
	US	EU	Commit	Defect
Defect	Commit		2,3,4	1,2,1
	Defect		1,2,1	1,2,1

Trade2:

JAP	US	EU	Commit	Defect
Commit	Commit		1,5,1	5,3,3
	Defect		3,4,2	4,1,4
	US	EU	Commit	Defect
Defect	Commit		2,2,5	4,1,4
	Defect		4,1,4	4,1,4

The Nash equilibria are marked by coloured cells. In the "Trade1" case, we find, besides the usual all-defect outcome, a Nash equilibrium in which the USA and Japan commit; this is the only stable coalition of these three games (see Table 8).

Table 8: Stable Coalitions.

	NT	T1	T2
Internally stable coalitions	{US,EU}	{US,EU}, {US,JAP}	Ø
Externally stable coalitions	{EU,JAP}	{US,JAP}	{US,EU},{EU,JAP}
Stable coalitions (both internally and externally)	Ø	{US,JAP}	Ø
Profitable coalitions	{US,EU}	{US,EU}, {US,JAP}, {US,EU,JAP}	Ø

Note : Stability and profitability based on the theory of cartel stability (see Section 4).

As mentioned before, we can identify 5 positions in this situation. Table 9 sums up all possible positions.

Table 9 Positions in the trade situation.

G^0	({1,2,3},{CCC},u)
G^1	({1,2,3},{DCC,CCC},u)
G^2	({1,2,3},{CDC,CCC},u)
G^3	({1,2,3},{CCD,CCC},u)
G^{123}	({1,2,3},{DDD, CCD,CDC,DCC,CCC},u)

Figure 3 shows how the inducement correspondence has to be reduced to account for coalitional moves.

Figure 3 The coalitional incremental commitment situation for three regions.

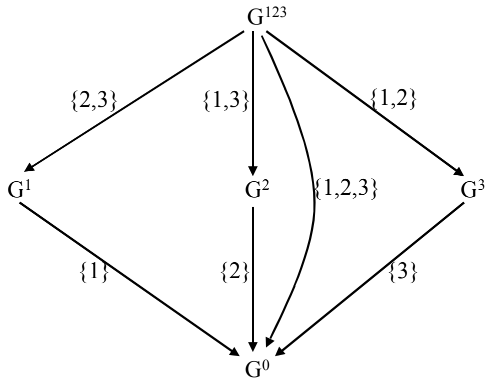


Table 10 shows the inducement correspondence in the specification of the trade game.

Table 10 The inducement correspondence in the trade situation.

$\gamma_{\{1,2,3\}}(G^{123}, \text{DDD}) = \gamma_{\{1\}}(G^1, \text{DCC}) = \gamma_{\{2\}}(G^2, \text{CDC}) = \gamma_{\{3\}}(G^3, \text{CCD}) = G^0$
$\gamma_{\{1,2\}}(G^{123}, \text{DDD}) = G^3$; $\gamma_{\{1,3\}}(G^{123}, \text{DDD}) = G^2$; $\gamma_{\{2,3\}}(G^{123}, \text{DDD}) = G^1$
$\gamma_s(G,x) = \emptyset$, otherwise for $G \in \Gamma$; $x \in X^G$

Table 11 shows the outcomes for the trade game.

Table 11 The ODOM for the trade situation.

No trade: ODOM(σ,G^{123}) = {DDD,CCD,CDC}	$\sigma(G^{123})$={EU,JP},{US,EU,JP}
Trade1: ODOM(σ,G^{123}) = {DDD,CCD,DCC}	$\sigma(G^{123})$={US,JP},{US,EU,JP}
Trade2: ODOM(σ,G^{123}) = {CDC}	$\sigma(G^{123})$=Ø, {US,JP}, {EU,JP}, {US,EU,JP}

The externally stable outcomes are outcomes in the solutions set. In the Trade2, situation, no agreement at all is an outcome in the solution set. This is because of the fact that as soon as a coalition is formed in the Trade2 situation, at least one member of the coalition is worse off. According to OSSB, this outcome may not be excluded from the solution set. This distinction was not possible based on the cartel stability argument. However, the grand coalition is also an outcome in the solution set.

6. Conclusions and recommendations

This paper models climate change negotiations as a social situation. We study three particular situations, namely the cartel situation, the incremental commitment situation and the trade situation. A clear advantage of using the theory of social situations (TOSS) instead of a conventional game theoretic model is that we have a tool for analysing off-equilibrium behaviour. While modelling the international negotiations for climate change as an n-person prisoner's dilemma, we cannot rule out some cooperation among world regions. Moreover, in the coalitional incremental commitment situation with three world regions, coalitional moves show that the grand coalition is an outcome in the solution set.

The purpose of this paper is to show the usefulness of modelling climate change negotiations as a social situation. However, the three situations in this paper have been rather restricted and kept as simple as possible to demonstrate how, computationally, to deal with the theory of social situations. Accordingly, none of the results is new. Carraro (2000) shows that small coalitions are formed in a game of international externalities with a large number of players. Eyckmans shows that myopic behaviour restricts the size of the maximum possible coalition. Kemfert *et al.* (2001) show that incentives to cooperate change if trade interactions are introduced. We reproduced these results, and put them in a unified framework, that allows a broader interpretation.

The areas for further research are manifold. First of all, the studied situations could be applied numerically to integrated assessment models on climate change to verify which regions of the world can form a coalition. As this paper shows, the size of the coalition can be larger than in the cartel game. Secondly, it is important to test empirically to what extent the suggested institutional rules correspond to the actual climate change negotiation process. A third area of further research, which is to study other situations, such as the individual and coalitional contingent threat situation. Finally, it is important from a theoretical point of view to guarantee for each situation that a stable standard of behaviour exists and that it is unique.

This paper shows an alternative way out of the prisoner's dilemma, namely by focusing on alternative behaviours, like farsightedness, commitment and coalitional moves. It would be useful to expand this model to see the impact of transfers and issue linkage.

From a policy perspective, we see that cooperative behaviour can emerge. This means that self-enforcing agreements are possible, provided that appropriate global institutions emerge.

References

Barrett, S., 1994, Self-Enforcing International Environmental Agreements, *Oxford Economic Papers* 46:878–894.

Barrett, S., 1998a, A theory of international cooperation. (UnPub)

Barrett, S., 1998b, The Credibility of Trade Sanctions in International Environmental Agreements. (UnPub)

Barrett, S., 1998c, On the theory and diplomacy of environmental treaty-making, *Environmental and Resource Economics* 11:317–333.

Botteon, M. and Carraro, C. (1997) Burden sharing and coalition stability in environmental negotiations with assymetric countries. In: Carraro, C., (Ed.) International environmental negotiations, strategic policy issues, pp. 26-55. Cheltenham: Edward Elgar.

Botteon, M. and Carraro, C. (1998) Strategies for environmental negotiations: issue linkage with heterogeneous countries. In: Hanley, N. and Folmer, H., (Eds.) *Game theory and the environment*, pp. 181-203. Cheltenham: Edward Elgar.

Carraro, C. and Siniscalco, D., 1993, Strategies for the International Protection of the Environment, *Journal of Public Economics* 52: 309–328.

Carraro, C. and Moriconi, F., 1997, International games on climate change control. (UnPub)

Carraro, C., 1997a, The Structure of International Agreements on Climate Change. (UnPub)

Carraro, C., 1997b, Environmental Conflict, Bargaining and Cooperation. (UnPub)

Carraro, C. (Ed), 1997c, *International Environmental Negotiations, Strategic Policy Issues*. Cheltenham, Glos, Edward Elgar,

Carraro, C., 1998, New economic theories. Impact on environmental economics, *Environmental and Resource Economics* 11:365–381.

Carraro, C., 2000, Coalition formation and EU leadership (UnPub).

Cesar,H., 1994. Control and Game Models of the Greenhouse Effect. Springer-Verlag, Berlin.

Chwe, M.S.Y., 1994, Farsighted Coalitional Stability, *Journal of Economic Theory* 63:299–325.

D'Aspremont, C. and Gabszewicz, J.J., 1986, On the stability of collusion, in Stiglitz, J.E. and Mathewson, G.F. (Eds.), *New Developments in the Analysis of Market Structure, Proceedings of a conference held by the International Economic Association in Ottawa, Canada*, London, MacMillan Press.

DeMarzo, P.M., 1992, Coalition, leadership and social norms: the power of suggestion in games, *Games and Economic Behaviour* 4:72–100.

Ecchia, G. and Mariotti, M., 1998, Coalition formation in international environmental agreements and the role of institutions, *European Economic Review* 42:573–582.

Eyckmans, J. 2000, On the farsighted stability of the Kyoto Protocol: Some simulation results (UnPub).

Fudenberg, D. and Masking, E., 1986, The Folk theorem in repeated games with discounting or with incomplete information, *Econometrica* 54:533–554.

Greenberg, J., 1990, *The Theory of Social Situations, an Alternative Game-theoretic Approach*. Cambridge, Cambridge University press,

Greenberg, J., Monderer, D., and Shitovitz, B., 1996, Multistage situations, *Econometrica* 64:1415–1437.

Hanley, N. and Folmer, H. (Eds), 1998, *Game Theory and the Environment*. Cheltenham, Edward Elgar.

IPCC, 2000, Presentation of Robert T. Watson, Chair, Intergovernmental Panel on Climate Change at the Sixth Conference of Parties to the United Nations Framework Convention on Climate Change, November 13, 2000, http://www.ipcc.ch/press/sp-cop6.htm.

Kemfert, C., Lise, W. and Tol, R.S.J., 2001, Games of climate change with international trade, Research Unit Sustainability and Global Change SGC-7, Centre for Marine and Climate Research, Hamburg University, Hamburg.

Mariotti, M., 1997, A model of agreements in strategic form games, *Journal of Economic Theory* 74:196–217.

Na, S. and Shin, H.S., 1998, International environmental agreements under uncertainty, *Oxford Economic Papers* 50:173–185.

Nakanishi, N. 1999. Reexamination of the international export quota game through the theory of social situations. *Games and Economic Behaviour* 27:132–152.

Pearce, D.W., Cline, W.R., Achanta, A.N., Fankhauser, S., Pachauri, R.K., Tol, R.S.J. and Vellinga, P., 1996, 'The Social Costs of Climate Change: Greenhouse Damage and the Benefits of Control', in Bruce, J.P., Lee, H. and Haites, E.F. (eds.) *Climate Change 1995: Economic and Social Dimensions -- Contribution of Working Group III to the Second Assessment Report of the Intergovernmental Panel on Climate Change*, Cambridge University Press, Cambridge. UK.

Ray, D. and Vohra, R., 1997, Equilibrium binding agreements: *Journal of Economic Theory* 73:30–78.

Tol, R.S.J., 1999, 'Kyoto, Efficiency, and Cost-Effectiveness: Applications of FUND', *Energy Journal Special Issue on the Costs of the Kyoto Protocol: A Multi-Model Evaluation*:130–156.

Smith, J.B., H.-J. Schellnhuber, M.Q. Mirza and others, 2001, 'Synthesis', Third Assessment Report of Working Group 3 of the Intergovernmental Panel on Climate Change, Cambridge University Press, Cambridge, UK.

Ulph and Maddison, 1997, Uncertainty, Learning and International Environmental Policy Coordination. *Environmental and Resource Economics* 9, 451–466.

Watson, R.T., Zinyowera, M.C. and Moss, R.H., 1996, Climate Change 1995: Impacts, Adaptation, and Mitigation of Climate Change –Scientific-Technical Analysis– Contribution of Working Group II to the Second Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge, Cambridge University Press. UK.